

## South American science at a crossroads

**Researchers in South America have a unique opportunity to advance their position in world science, provided they champion necessary reforms of universities and research agencies.**

**T**he outlook for South America's political and intellectual life offers more genuine grounds for optimism than has been the case for decades. Democracy is more reliably ensconced in Argentina, Brazil, and Chile than ever before, while low inflation and trade liberalization are transforming the continent's economy. All this presents a unique and unrepeatable opportunity for the region to attain its rightful place in the world of science.

As Fernando Reinach of the University of São Paulo explains in this issue (pp 647–648), the opportunity involves sizeable risks for the region's large and talented scientific community. Through decades of political instability, protectionism, occasional hyperinflation and frequent military dictatorship, the situation of scientists and other academics has been ambiguous at best. Nationalist governments have made reasonably strong investments in research and development while accepting low standards of performance. Academics have taken refuge behind constitutional protection of intellectual freedom, fiercely resisting outside influence.

It was understandable during much of the recent past that academics viewed tenure and a degree of self-government as more important than instilling competition in the universities. But the result today is a university sector with no external oversight and little incentive to excel in research. As Reinach points out, this sector is not immediately well-placed to support the science and technology requirements of the region's new economy, in which old state-run industries have been exposed to domestic and foreign competition.

The same applies to the disproportionately large number of scientists in the region who work for government research laboratories. The national council for science and technology (CONICET) in Argentina, for example, spends nearly all of its money on 3,000 staff scientists at its own institutes, leaving only a paltry amount for exter-

nal grants. The Argentinian government is justly suspicious of CONICET's claim that all its work is being properly peer-reviewed (see *Nature* 391, 525; 1998).

Universities have poorly developed ties with the continent's newly energized industrial base, which may go outside for technical help. Government-funded scientists often work at universities but do no teaching. These gaps between science, education and commerce must be closed. Otherwise there is a serious risk that the universities and the scientists will slip into irrelevance.

Moves are already under way in the region that will modernize and strengthen its science base. A new National Agency for the Promotion of Science and Technology has been established in Argentina, for example, to distribute extramural research grants. The Inter-American Development Bank played a significant role in this initiative — as did the World Bank in a comparable project in Brazil (see *Nature* 391, 317; 1998). An important objective of both exercises is to develop a structure for rigorously competitive peer review.

David Sabatini, an Argentinian who chairs the department of cell biology at New York University, has meanwhile come up with an imaginative proposal for a chain of regional centres of excellence in the life sciences, sharing the resources and talent of Argentina, Brazil, Chile and possibly Mexico as well. Fifty years of European experience suggests that such joint ventures between nations can be an effective means of generating political cohesion, as well as scientific excellence.

Such initiatives can pay generous dividends in a continent blessed with a large and talented scientific community which, although largely educated in the United States and in Europe, is strongly inclined to return home to do science. That community must now lead extensive reform of the region's universities and scientific institutions, rather than entrench themselves against it. □

## Mirage in the Senate?

**Enthusiasm for science is strong in Washington, but enthusiasm for necessary choices is not.**

**T**his month's US Senate budget resolution is the best pointer yet to what next year's science budget will contain, and the signals it sends are mixed. On the one hand, the resolution specifies a generous increase for research at the National Institutes of Health. On the other hand, it does not do much for research at other agencies.

The budget resolution groups together the National Aeronautics and Space Administration, the National Science Foundation and the civilian research activities of the Department of Energy in a category called 'general science, space and technology'. The Senate resolution offers a small increase to this element of the budget next year, followed by steady decline for the following four years. But it includes an amendment stating that it is 'the sense of the Senate' that spending on scientific research should double over the next ten years. The science community can thank Representative George Brown (Democrat, California) for pointing out this incongruity (see page 637).

To be fair, President Bill Clinton did no better in his February budget, which boosted science and other favoured programmes with money which he did not save elsewhere. The Senate has now entered into the same spirit, unanimously passing a non-binding motion to double spending on scientific research in a budget resolution that will, in fact, leave appropriators hard pushed to come up with any increases at all for key science agencies.

The best hope for bridging the gap between rhetoric and budget lines is still the so-called 'tobacco settlement', which each day looks less like a public health measure and more like a tool to satisfy Washington's spending addiction by means of new and regressive taxation. It is regrettable that neither the administration nor the Senate has had the courage to meet their stated spending priorities from within the limits to which they themselves agreed under last year's balanced budget agreement. □



# US legislator warns of hollow promises on research spending

[WASHINGTON] US scientists are fooling themselves if they expect the Congress to honour recent pledges to double research spending over the next decade, according to one of the country's most experienced science legislators.

An analysis conducted for the office of George Brown (Democrat, California) by the Congressional Research Service says that the budget resolution passed by the Senate on 2 April would actually result in a small decline in spending on civilian research and development over the next five years.

Spending in agencies other than the National Institutes of Health (NIH) would drop by 17 per cent, from \$22 billion to \$18.5 billion by the year 2003.

NIH would receive an 11 per cent increase in funding next year, to \$15.1 billion, under the budget resolution, an important guide to the budget for the 1999 fiscal year which Congress must agree on by 1 October.

Brown is the senior Democrat on the Science Committee in the House of Representatives. His committee oversees research at the National Science Foundation (NSF), the National Aeronautics and Space Administration and the Department of Energy, and his attack on the resolution reflects his concern that funding prospects are darkening for science agencies other than NIH.

Brown argues that it is the budget resolution, rather than a separate bill proposed in the Senate that would double research spending over ten years, that shows the real intentions towards science of senior Republican senators.

Two sponsors of the bill, known as S1305 — Senators Phil Gramm (Republican, Texas) and Pete Domenici (Republican, New Mexico) — sit on the budget committee, which drew up the resolution. Brown says: "Even with key champions of S1305 in the room, when it came to making hard choices, the cheap talk stopped and dollars went to other purposes."

Larry Neill, a spokesman for Gramm, responds to the charge by accusing Brown of "indulging in cheap partisanship" which could endanger the S1305 proposal. Neill says that Gramm hoped to pass the bill this year, with a view to influencing budget resolutions in future years. But there is no companion measure in the House of Representatives, and it is not considered likely to pass into law.

Gramm did successfully propose a 'sense of the Senate' amendment to the budget resolution, stating that the Senate intends

civilian research and development spending to be doubled over ten years. But the Senate passed several such amendments to its budget resolution, and critics point out that they allow the legislative body to commit words without spending any money.

Prospects for research increases in the budget for 1999 remain uncertain. Despite doubts about the achievement of a tobacco settlement that would finance such increases in the budget submitted to the Congress by President Bill Clinton in February (see *Nature* 391, 521–522; 1998), most science lobbyists are optimistic that money will be found to pay for the proposed increases.

The appropriations subcommittees that work on the budget have begun their annual process of reviewing the science agencies, although they do not yet know how much

money they will have to fund the programmes they oversee. The House has not yet drawn up a budget resolution to match the Senate's, and is not expected to do so until next month.

According to science lobbyists, the appropriations subcommittees are likely to receive initial funding allocations in late spring that will be insufficient to meet their members' aspirations. Attempts will then be made to find extra money — from a tobacco settlement or some other source.

A last-minute scramble is then likely to take place in September to ensure that key programmes are funded before congressmen return triumphantly to their districts to seek approval in November's elections from the people whose money they have spent.

Colin Macilwain

## Europe may pool marine research efforts

[PARIS] Three marine research agencies in Europe are planning to take further steps towards the creation of a unified European fleet of research vessels with a common scientific programme. The initiative represents a major step towards integrating the agencies' operations, and coincides with separate preliminary plans to float a European Maritime Agency (see *Nature* 392, 323; 1998).

The lead in the project is being taken by the Triangular Liaison Group (G3), which was set up by the principal bodies responsible for marine research in the United Kingdom, France and Germany in 1996 to strengthen cooperation on large marine facilities. The group has already agreed with the ship operators of each agency on pooling the use of their research vessels (see *Nature* 379, 576; 1996).

G3 brings together Britain's Natural Environment Research Council (NERC), the French Institute for Research into the Exploitation of the Sea (IFREMER), and Germany's ministry of research, education and technology (BMBF).

Concrete steps towards closer integration are expected to emerge from the group's next meeting in June. This will bring together not only the research ship operators, but also representatives of the peer review committees that approve research strategy within each agency.

According to Pierre Nounou, the official responsible for the G3 at IFREMER, the initial goal is to agree on common research programmes, with the ultimate aim of having a single evaluation committee for



Navigating towards unification: vessels such as *Cyana* may come together for European research.

research proposals and a single group of ship operators. "This would amount to a European fleet of research vessels," he says, adding that the construction of new joint European vessels is also on the agenda.

The current agreement is limited to Germany, France and the United Kingdom, which together represent more than 80 per cent of Europe's marine science resources. These countries are also the world's foremost marine science powers with a combined marine research workforce of more than 7,000 and a budget of FF4.3 billion (US\$700 million) — roughly equivalent to that of the United States, Canada and Japan combined.

Under longer-term plans, the tripartite agreement would be opened to the smaller European countries. These have a joint research workforce of 1,420 and a budget of FF930 million, but often lack ships, submarines and other large facilities. Researchers from these countries could be included in research projects, for example, in return for a subscription.

Declan Butler

## Russian acoustics scientists defended on 'secrets' charges

[MOSCOW] Dozens of prominent Russian scientists have written to the Federal Security Service in support of their colleagues Vladimir Borodin and Mikail Galaktionov, of the Institute of Acoustics in Moscow, who could be accused of passing state secrets to a foreign company. The scientists claim the information had already been published.

Security service officers searched the institute laboratory headed by Borodin and found scientific reports prepared for the US Lockheed Sanders Corporation that had not been authorized by the institute. Nikolai Dubrovsky, the institute's director, launched an inquiry that concluded that the reports should not have been passed to the company.

But Borodin, who was not invited to participate in the inquiry, says that details of his research on which the reports were based had been published in the institute's proceedings in 1995.

The Institute of Acoustics used to be one of the most prosperous research centres in Russia, generously financed by the military. But in 1994 it suddenly found itself on the verge of poverty. Borodin tried to save his department by entering a contest announced by Lockheed Sanders for Russian scientists. They were invited to collaborate in developing a system to protect oil derricks in shallow waters from damage by submarines.

Dubrovsky prohibited Borodin's lab from taking part in the contest, but Borodin and Galaktionov decided to enter as private individuals. They won the competition and started the research. Later, when they needed money for experiments but could not guarantee successful results, Lockheed terminated the contract.

Dubrovsky has not explained why the work is considered secret. But A. Luchinin, a specialist in hydro-acoustic protection systems who is not connected with the acoustics institute, told the inquiry that the work does not contain secrets. The same conclusion was reached at a seminar held by the Russian Academy of Sciences last year.

Borodin says: "The worst thing that could happen is that Galaktionov will leave science. He is now having to work for a commercial firm. Galaktionov's institute salary is 255 rubles (US\$42) a month, but he needs to pay for expensive medical treatment for his child, who underwent a complex operation in France when Galaktionov worked there."

The operation was done free of charge after his French colleagues appealed to the then President François Mitterrand. "But now Galaktionov is in Russia, where the authorities are concerned about keeping secrets, not people," says Borodin.

Carl Levitin

## Fight heats up over Toronto racial discrimination claim

[MONTREAL] The University of Toronto is coming under renewed pressure to resolve a long-standing battle with a seismologist of Chinese descent who has accused the university of racial discrimination by repeatedly denying him a tenure-track professorship.

Kin-Yip Chun joined the university's physics department in 1985, and received no salary during almost 10 years of service, deriving his income entirely from external research contracts. During that time, however, he taught students, published 26 papers, and attracted more than C\$1.4 million (US\$1 million) in research grants.

Cecil Yip, former vice-dean of the medical faculty, was appointed by the university to investigate Chun's charges that he had been passed over for tenure. In 1994, Yip concluded that, although he could find no evidence of discrimination on the basis of race, Chun had been "exploited" by the physics department. Shortly afterwards, Chun was escorted off the university site by campus police, and banned from it for three years. No reason was given at the time.

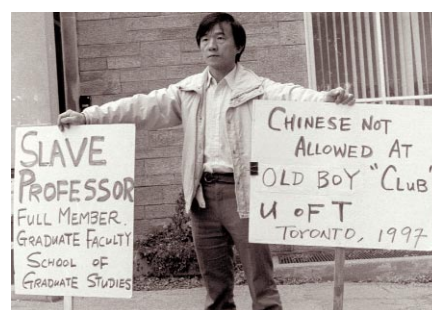
The academic freedom and tenure committee of the Canadian Association of University Teachers (CAUT) has expressed concern about the university's failure to resolve the issue. The committee recently informed the university president, Robert Prichard, that it is preparing a report on the case.

Prichard has been invited to see the report before publication — which may be as early as next month — and to publish a commentary alongside it. At a meeting of the university's academic board last year, at which Chun's case was discussed, Prichard argued that the university had done nothing wrong.

The university's faculty association has complained of "unfair treatment" to Chun, who has received support from faculty members, students and others. The Ontario Human Rights Commission has appointed an investigator to look into the case.

Paul Gooch, the university's vice-provost for staff relations, says Chun did not win tenure-track competitions because better candidates were found. He says repeated offers of employment were made to Chun, including offers to place him elsewhere.

But Chun argues that those who won tenure were less qualified than he was, and that the university's offers would give him only the same status as his former assistants, of whom he had the largest team in geophysics in Canada. Chun adds that, unlike other university professors — and contrary to the rules of the principal federal grants bodies — he would still have to find his salary out of grants or contracts. He says that,



Chun: claims he had to carry out a professor's duties while working as a research associate.

although the university promised to try to find positions for him with other employers, they never came up with one.

Chun graduated from the University of Toronto with a bachelor's degree in engineering science, and won a master's from Columbia University in solid Earth physics and a PhD from the University of California, Berkeley. After postdoc studies at Berkeley, he was attracted back to Toronto by an offer of a post as research associate. But on arrival he was told there were no funds to pay him.

Chun says that he was obliged to work as if he were a professor. He claims the university arranged for him to apply for external research funds, and that within six months he had attracted the largest grant in the history of the geophysics division.

The CAUT says Chun was apparently subject "to unfair procedures during at least two of the competitions for tenure-track positions in the physics department". The cancellation of a competition in 1987 in mid-stream to reappoint a previous faculty member without competition denied him fair consideration for the post, says the association. Yip's report suggested that this could have been done to keep Chun out of the job.

A Web site for Chun's support committee (<http://www.utoronto.ca/acc/chun>) says the university has never appointed a tenure or tenure-track geology or geophysics professor from a minority group. Three years ago, the CAUT claimed that "there are aspects of [Chun's] case which arouse a suspicion of [such] discrimination".

Derek Paul, physics professor emeritus at the university, wrote that, although the Yip report failed to uncover "overt racial bias", if it had probed more deeply it would probably have found "systemic" racial bias in the university hiring, and have refused to rule this out in Chun's case. "A lot of alumni are very concerned about the university, whose own investigation says it's exploited somebody," says Patrick O'Neill, chair of the CAUT committee.

David Spurgeon



## US space agency sharpens focus on near-Earth objects

[WASHINGTON] The US space agency NASA plans to double its budget for studying near-Earth objects this year, following a meeting last month to assess the progress of current asteroid searches (see *Nature* 392, 215; 1998).

The increase to \$3 million affects only the 1998 budget, says NASA's Solar System exploration chief, Carl Pilcher. Future spending is yet to be determined. The agency will also set up a new office to coordinate research on near-Earth objects.

NASA underwrites the lion's share of such research in the United States, with additional funding coming from the air force. But past spending levels — between \$1 million and \$1.5 million a year — are insufficient to achieve the long-held goal of identifying most large, threatening asteroids within a decade. An estimated 90 per cent of such objects have yet to be detected.

The new NASA office will decide how to augment several searches for near-Earth objects already under way or planned. It will also decide policy for informing the public about future asteroid threats, which many scientists believe was bungled last month. Under a draft policy proposed last week, NASA-funded researchers would have to notify the agency and consult other asteroid scientists before issuing public statements.

Data on near-Earth objects received at the Minor Planet Center (MPC) at the Harvard-Smithsonian Center for Astrophysics in Cambridge, Massachusetts, would be released to scientists "generally within 24 hours of receipt". Asteroid scientists have accused the MPC of not sharing data promptly.

Some non-US asteroid researchers are uneasy about a US government agency dictating who should talk to whom when an Earth-threatening object is discovered. The MPC, they say, operates under the auspices of the International Astronomical Union (IAU), and should remain independent.

But IAU's assistant general secretary, Hans Rickman of the Uppsala Astronomical Observatory in Sweden, sees no problem, as long as NASA policy is not presumed to apply to all asteroid researchers worldwide. He says NASA "has every right to require certain things" of the researchers it funds. The IAU should formulate its own policy on asteroid warnings and data sharing, he says.

Rickman agrees that dissemination of data by the MPC "could be done more efficiently". But he points out that the centre, which serves the entire international astronomical community, currently operates under severe financial and staffing constraints.

**Tony Reichhardt**

## Petition strengthens hand of global warming sceptics

[WASHINGTON] About 15,000 US science graduates, including 6,000 PhDs, have signed a petition that rejects the Kyoto agreement on global warming and argues that increases in carbon dioxide levels benefit Earth, according to the petition's organizers.

The petition is likely to be released by the George Marshall Institute in Washington DC in the near future, and opponents of the Kyoto agreement are expected to use it to back up their arguments about a lack of scientific consensus on the issue.

But the mass-mailing of the petition to scientists — accompanied by a lengthy review article that had not been peer-reviewed or published — has angered some of those who believe that carbon dioxide emissions are a serious problem. They also argue that the number of signatories is a relatively small proportion of those who were mailed.

"Virtually every scientist in every field got it," says Robert Park, a professor of physics at the University of Maryland at College Park and spokesman for the American Physical Society. "That's a big mailing." According to the National Science Foundation, there are more than half a million science or engineering PhDs in the United States, and ten million individuals with first degrees in science or engineering.

Arthur Robinson, president of the Oregon Institute of Science and Medicine, the small, privately funded institute that circulated the petition, declines to say how many copies were sent out. "We're not willing to

have our opponents attack us with that number, and say that the rest of the recipients are against us," he says, adding that the response was "outstanding" for a direct mail shot.

The Union of Concerned Scientists has branded the exercise as "a deliberate attempt to deceive the scientific community with misinformation on the subject of climate change". And prominent members of the National Academy of Sciences, whose past-president, Frederick Seitz, wrote a cover letter for the mailing, are also upset, according to a spokesperson for the academy, partly because the article in the mailing looks exactly like a paper from the *Proceedings of the National Academy of Sciences (PNAS)*.

Robinson, a biochemist and former close associate of Linus Pauling, co-authored the article with Sallie Baliunas, a planetary scientist at Harvard University and noted global warming sceptic, and with Robinson's son Zachary, who has just obtained his bachelor's degree from Oregon State University. The article, "Environmental effects of increased atmospheric carbon dioxide", argues that the release of more carbon dioxide "will help to maintain and improve the health, longevity, prosperity and productivity of all people".

"I really don't think we are deceiving people," says Robinson. "I have published in the past in *PNAS* and I chose the format because I liked it. I wanted the format that scientists are used to reading." Robinson says the paper is now being submitted to an undisclosed journal for publication.

**Colin Macilwain**

## Teachers gain aid for evolutionary struggle

[WASHINGTON] The National Academy of Sciences, alarmed by what it says is continuing hostility to the teaching of evolution in many school districts in the United States, has produced a glossy guide to advise teachers on how the subject can be best taught.

The guide is designed to help teachers in parts of the country, primarily in the south, where Christian groups have tried to ensure that 'creation science' is taught alongside evolution. It says that the Supreme Court rejected that idea in 1987, when it held that a Louisiana decree calling for the "balanced treatment" of the two was unconstitutional.

"We're not saying that teachers can flaunt their state rules," says Maxine Singer, president of the Carnegie Institution of Washington and one of the guide's authors. "But they could challenge them legally" on the basis of the Supreme Court ruling, she says.

The guide says that the emergence of genetics has made the theory of evolution a central tenet of biology, which teachers can-



**Singer: option of legal challenge.**

not choose to ignore. Many teachers mistakenly assume that evolution has little to do with science as it is practised today, says Singer. "A few years ago, evolution was not of major interest to molecular biologists," she says. "But now it is."

According to academy officials, a few states have education policies that require the teaching of creationism. Schools are usually administered, however, by counties and problems arise in all regions of the United States.

Donald Kennedy of Stanford University, who chaired the group of 13 authors, declines to estimate how many teachers are intimidated by such policies.

"If you listen to enough teachers you become persuaded that this is a serious problem," Kennedy says.

**C.M.**

# UN eco-fund under pressure to open up

[NEW DELHI & LONDON] The managers of the main United Nations (UN) fund for environmental research projects have agreed to review their policy on supporting projects in developing countries. The move follows criticism of the fund's excessive bureaucracy and apparent insensitivity to local concerns.

The Global Environment Facility (GEF), a US\$2 billion UN fund organized through the World Bank, concluded at its first assembly in New Delhi, India, two weeks ago that its activities need to be more transparent, and agreed to accept greater involvement from the private sector, the public and environmentalist groups.

But the meeting failed to resolve a long-running controversy about the method used to cost projects, agreeing simply to show "greater flexibility" in assessing applications.

There was controversy also about several other issues. These included complaints that the fund is too small, that there is insufficient transparency in assessing applications, that public involvement is cursory and

that the categories of projects eligible for funding reflect the concerns of developed, rather than developing, countries.

India, Brazil and several African countries, for example, had pressed for additional GEF funds to go to projects to improve water supply and sanitation, or to slow down desertification — both prime concerns for the African states. They argued that such issues are of more pressing environmental concern than climate change or ozone depletion.

But GEF managers maintained that there would be no change in distribution of GEF funds to the four 'focal areas': climate change and biodiversity conservation (40 per cent each), and ozone depletion and water supplies (10 per cent each).

The findings of a 26-country survey of GEF-fund recipients fuelled much of the debate. The survey, carried out by the secretariat of the UN Convention on Biological Diversity, revealed concern about the scale of spending on international consultants,

whose costs often consume up to half the total project expenditure. It also revealed deep unease over the method used to calculate the size of project grants.

Under GEF rules, a country must show it can finance a project in full, and the GEF then pays for the extra cost of a more environmentally friendly alternative. This system, known as 'incremental costs', was developed after the Second World War to help finance the reconstruction of Europe.

But the countries replying to the survey complained that incremental costs were unworkable for activities such as research, where the concept of a more environmentally friendly alternative does not hold. According to the survey, "the principle of incremental costs in the case of biodiversity was meaningless, and decisions in this respect were arbitrary and confusing".

The survey also reflected concern in some countries that projects take too long to get approval. Suresh Praby, India's environment minister, said most of India's 12 GEF projects have yet to take off. He claimed that proposals have to be sent to the GEF council at least three times before they are approved, and that the total process takes about three years.

But a GEF spokesman in Washington says projects are usually approved within three months — less time than the World Bank's non-GEF activities take.

Meanwhile, the announcement at the New Delhi meeting of a \$2.75 billion, three-year replenishment of the GEF fund was marred by complaints that it contained \$680 million brought forward from the first round of GEF funding, as well as \$80 million of unused funds. Moreover, contributions from the United States, Germany and Italy were lower than for the first round.

In addition, some emphasized that the fund is considerably smaller than other forms of development assistance for similar projects. Halifax Initiative, for example, a Canadian environmentalist group, pointed out that the World Bank's support for fossil fuel projects undermines GEF's efforts to support cleaner technologies.

GEF's own analysis shows that between 1993 and 1997, the bank committed \$9.4 billion for carbon-dioxide-emitting fossil fuel projects, while setting aside a relatively meagre \$500 million for GEF projects.

But Caio Koch-Weser, managing director at the World Bank, says the bank is preparing to devote more non-GEF finance to environment-friendly projects. These will include a scheme for trading carbon emissions (see *Nature* 390, 7; 1997), and a \$186 million wind power scheme enabling India to raise its wind power generation from 30 to 700 MW.

**K.S. Jayaraman & Ehsan Masood**

## Ministry 'changing course' on lab closure

[LONDON] Britain's Ministry of Agriculture, Fisheries and Food (MAFF), currently planning to close its food research laboratory in Norwich, decided 16 months ago that such a move would be detrimental to its scientific activities, according to an internal document released by a member of parliament.

That decision, concerning the Norwich branch of MAFF's Central Science Laboratory, was made in December 1996 following the 'prior options' review of the government's ownership and management of its laboratories.

A document dated February 1997 describing the decision was released this month by Charles Clarke, Labour Member of Parliament for Norwich South as part of a campaign to thwart renewed plans to close the Norwich laboratory and 'rationalize' MAFF's food research at a single complex in York (see *Nature* 392, 319; 1998).

Clarke argues that the government will be contradicting its own conclusions if it goes ahead with the closure. A MAFF spokesman declines to comment on the statement as the document was written "before our time". But he says that no formal decision has been taken on whether the laboratory will close.

Clarke quotes from the statement in an 18-page letter to the agriculture secretary, Jack Cunningham, urging him not to relocate the 137 staff at the laboratory to a new £133 million (US\$222 million) research complex in York.

Clarke has asked more than 50 questions in parliament in the past few weeks, partly

in an attempt to draw public and political attention to the issue.

Richard Packer, MAFF's permanent secretary — its senior civil servant — is proposing the move to help fill unused space at York, as the ministry has to pay for unused space under a new government accounting system. In his letter, Clarke says Packer's advice appears rooted in "a desire to justify retrospectively the decision to build the [York] laboratories, rather than a desire to increase the quality and integrity of food research".

The 1997 statement says that, despite potential cost savings, the laboratory should stay in Norwich to retain its research contracts and sustain its cooperation with the Institute of Food Research, operated by the Biotechnology and Biological Sciences Research Council. Both laboratories are based in the Norwich Research Park, along with the John Innes Centre.

The Norwich food research laboratories account jointly for around one-third of Britain's food research funding — three times more than the next largest recipient. According to unofficial estimates, fewer than half of Norwich's 109 scientists are likely to relocate to York should the move take place.

But the proposed move has put the main researchers' trade union, the Institute for Professionals, Managers and Scientists, in a difficult position. The institute represents scientists at both York and Norwich. One union member says it intends to keep a low profile as "we do not want to pitch scientist against scientist".

**E.M.**

## Curbs on research dropped from US medical privacy bill

[WASHINGTON] A Republican senator has introduced in the US Congress a medical privacy bill that essentially preserves existing procedures for researchers seeking access to identifiable patient information. This is in contrast to a draft of the bill that would have imposed heavy restrictions on researchers and industry (see *Nature* 392, 6; 1998).

The new version of the bill, from Senator James Jeffords, lacks any requirement for increased scrutiny of Institutional Review Board (IRB) waivers of informed consent. Such waivers are currently allowed for researchers seeking access to identifiable records in cases where it would be impractical to carry out the research if informed consent had to be obtained.

Nor does the Health Care Personal Information Nondisclosure Bill eliminate — as the draft did — ‘expedited review’ for such projects, a procedure under which IRBs may designate a member to give prompt approval.

The bill has also dropped an extension, to the private sector, of protection rules for human subjects, including informed consent for use of identifiable records, that now apply only to federally funded scientists. Industry had said the requirement would be crippling.

“We have worked hard with researchers to make sure we have provided them with the tools necessary” to their work, Jeffords said. “I believe that we have accommodated [their earlier] problems.”

Meredith Wadman

## CERN told to start technical thinking for next collider

[MUNICH] The European Laboratory for Particle Physics (CERN) in Geneva, Switzerland, has been urged to begin technical design work for the successor to its Large Hadron Collider (LHC), now being built and due to come into operation in 2005.

The SFr2.6 billion (US\$1.7 billion) LHC stretches the limits of accelerator technology, and has an expected lifetime of up to 20 years. Its high energy of 14 TeV should allow it to detect the Higgs boson and supersymmetry. An internal report points out that its successor would need to operate at even higher energies. This would require the development of more powerful accelerating systems, or longer tunnels — both very expensive options.

The report was commissioned last summer by CERN’s director general, Christopher Llewellyn Smith, and prepared by a small group of CERN particle physicists. It has now been presented to CERN’s science programme committee. It suggests that future design work should focus on the development of cost-effective high-field magnets.

It identifies three possibilities for future accelerator concepts. The most studied, in the United States and Japan as well as at CERN, is a 5 TeV electron collider. A second option, a muon collider, is particularly challenging technically because of the rapid rate of decay of muons.

A third option would be a Future Large Hadron Collider (FLHC), a design concept for which has already been developed in the United States (see *Nature* 385, 471; 1997). This would use a large-circumference, narrow tunnel, built with technology used for laying sewerage pipes. The narrow bore of the tunnel would require the ring to be controlled robotically, rather than by technicians. An electron-positron collider could conceivably share the same tunnel, adding to the possible scientific payoff.

The report says that it is important for CERN to study these concepts now because the technologies will take decades to develop. It will also be many years before particle physicists know what questions remain to be answered after the LHC and other relatively high-energy colliders now being planned have yielded their secrets, and which new type of collider could best answer them.

Tunnel requirements would influence the decision. CERN could conceivably build a tunnel for a muon accelerator at its site, as the characteristics of muons mean that it would be relatively compact. A 30 km tunnel suitable for a linear electron collider could be located along the nearby Jura mountain range. But the report admits that CERN might find it hard to find a site for a tunnel to meet the requirements of an FLHC, which would need to be at least 120 km in circumference.

Alison Abbott

## Collaboration is the name of the game for Europe’s scientists

[MUNICH] Cooperation between scientists in different European Union (EU) countries has increased significantly over the past decade, according to statistics released by the European Commission last week.

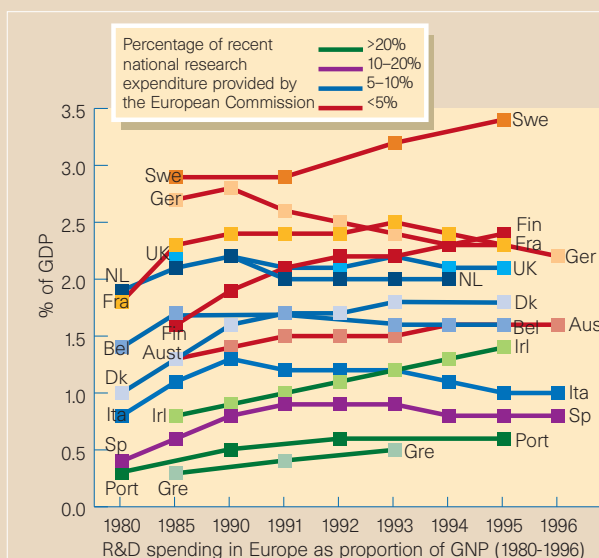
In the ten years from 1985, the funding of research and development in European collaborative programmes rose from 6 to 16 per cent of total spending. That includes the commission’s Framework programmes, Eureka and the European-level laboratories.

Framework programmes are designed to stimulate cooperation between researchers in different countries and between industry — particularly small and medium-sized enterprises — and academics. The report says that more than

200,000 such links were established between 1990 and 1996, and that 90 per cent of these were international.

Overall, the statistics show that Europe’s international research standing has changed little. The EU invests less in science and technology than its major competitors, the United States and Japan, as a percentage of gross domestic product (1.8, compared with 2.5 and 2.8 respectively). And the number of researchers in Europe is much lower — 5 per 1,000 of population. Scientific output, in terms of publications, remains equal to that of the United States, and above that of Japan.

European-level funding is becoming increasingly important for research. Although the proportion of national public



expenditure on research and development fell from an average of 3.5 per cent in 1985 to 2.6 per

cent in 1995, research's share of the commission's general budget rose from 2.1 to 5.1 per cent. **A.A.**



# Call to boost research on particulates

[WASHINGTON] The United States should launch a long-term research project into the effects on health of airborne particulate matter (PM), according to a committee of the National Research Council (NRC), part of the National Academy of Sciences complex. The committee recommends a 13-year programme costing \$440 million.

Such a programme, if approved, would be comparable in scope to a comprehensive decade-long acid rain study that was conducted in the 1980s.

The NRC committee, which began work in January and will continue for five years, was asked by Congress to set an agenda for PM research after the Environmental Protection Agency (EPA) enacted controversial new regulations governing particulate emissions last year (see *Nature* 388, 5; 1997).

Both sides in the bitter regulatory debate agreed that more research is needed. The panel chairman, Jonathan Samet, an epidemiologist at the Johns Hopkins University in Baltimore, Maryland, says the committee's first report, released last month, is the beginning of an attempt to develop a long-term, integrated research plan.

The committee outlined ten research areas it considers high priority for a PM science programme. EPA's current and planned activities are "generally reasonable and

potentially useful," it says. But the agency should pay more attention to investigating the relationships between fixed-site outdoor monitoring data and personal exposure to PM. And, rather than initiating epidemiological studies, EPA should concentrate on understanding the mechanisms behind the effects on health attributed to particulates.

The committee was careful not to take a stand on whether EPA's new rules governing particles smaller than 2.5 micrometres are justified on scientific grounds. But it did express concern that the agency is rushing to establish a nationwide network for monitoring PM<sub>2.5</sub> without fully considering the needs of researchers.

Monitors for establishing compliance with the law may not be sophisticated enough to characterize the size and chemistry of airborne particles — information required for a broader research programme.

John Bachmann, of EPA's Office of Air Quality Planning and Standards, says that the agency fully intends to use both kinds of monitors, and to consult scientists fully before putting the network in place. His office will hold a workshop next month to solicit the research community's views.

One panel member, Ronald White of the American Lung Association, says the committee simply wanted to emphasize the criti-

cal need for research-quality monitoring data. So far, he says, EPA appears to be emphasizing the regulatory role of the monitoring network, and the "health research community is feeling disenfranchised".

The NRC panel also encouraged EPA to accelerate plans to set up several independent university research centres for PM research. Congress designated \$8 million for this purpose in 1998. But it will take time, according to one EPA source, to review applicants and negotiate contracts. A request for proposals is expected to go out this month.

The NRC committee outlined a 13-year PM research programme ending in 2010, with annual average spending of more than \$50 million from 1999 to 2003. Congress almost doubled EPA's request for PM research last year, granting \$49 million for 1998. But the agency has asked for only \$29 million in 1999 — partly because it was waiting to see what the academy would recommend.

Opinions are divided on whether legislators will approve a long-term increase. One congressional staff member with oversight of EPA's budget says \$50 million a year is "not very realistic". But another congressional source says that even lawmakers who opposed the new EPA rules favour more research, as they think it will disprove the need for tighter regulation. **Tony Reichhardt**

## US and Japanese scientists in dispute over 'poisoned' radishes

[TOKYO] Scientists from the US Food and Drug Administration (FDA) are contesting a claim by Japan's health ministry that radish sprout seeds imported from the United States were to blame for last year's food poisoning outbreak caused by the O-157 strain of *Escherichia coli*. FDA scientists are seeking consultations with ministry officials over the issue.

A report from the ministry's Food Sanitation Investigation Council concludes that an outbreak of food poisoning last March in Aichi and Kanagawa prefectures, which affected 115 people, was caused by eating white radish sprouts grown from contaminated seeds imported from Oregon.

The report says the O-157 *E. coli* bacterium was not found in the seed sample. But the presence of genetic material linked to vero toxin — a byproduct left by the bacterium — as well as genetic fragments claimed to be unique to O-157 *E. coli*, were detected.

FDA scientists remain unconvinced, arguing that the test results "do not adequately support the finding that the seeds were contaminated with *E. coli* O-157". In a statement released by the US Embassy in Tokyo, the FDA points out that the



investigation did not involve a direct comparison of O-157 *E. coli* isolated from patients and samples extracted from the seeds, as would be necessary to conclude that *E. coli* from the seeds was the likely cause of food poisoning.

Explaining why no trace of the bacterium was detected in the seeds under conventional cultivation, ministry officials say that strains of *E. coli* other than O-157 proliferate much faster, suppressing the growth of O-157 colonies. Despite

announcing its intention to enforce mandatory sterilization of radish sprout seeds before cultivation, Japan's Ministry of Agriculture, Fisheries and Forestry is also sceptical about the health ministry's conclusion. "Vero toxins are also found in bacteria other than O-157 *E. coli*," says a spokesman for the agriculture ministry.

"The test results merely suggest the presence of the bacteria in seed samples; they do not verify the cause of the food poisoning outbreak."

Japanese producers of white radish sprouts say that the health ministry's recent announcement has dealt a new blow to their business. The ministry set off a public health scare in 1996 when it announced that white radish sprouts might be the source of an earlier O-157 food poisoning outbreak in the Osaka area, which affected more than 9,000 people and claimed nine lives (see *Nature* 382, 567; 1996).

The ministry subsequently came under fire for reaching a hasty and unsubstantiated conclusion. Its announcement forced food distributors to withdraw radish sprouts from shops and restaurants, and drove many farmers to bankruptcy. **Asako Saegusa**



## Early end to tamoxifen cancer trial leaves questions unanswered

[LONDON] British scientists expressed fears last week that the suspension of a US trial of the anti-breast cancer drug tamoxifen could wreck their own parallel seven-country trial. The US National Cancer Institute (NCI) halted its six-year trial of 13,000 women when researchers reported 85 cancers among patients taking tamoxifen, compared with 154 among those on a placebo.

The trial had 14 months left to run. British scientists fear that the parallel International Breast Cancer Intervention Study (IBIS) may collapse if its 7,000 participants on placebos insist on being given tamoxifen. One IBIS scientist says that, although interim findings show that tamoxifen appears to reduce the incidence of breast cancer, researchers may never know whether the drug can prevent the disease in the longer term. "We will not know how many of them will develop cancer in the years to come," he says. "My feeling is that, if we stop trials prematurely, we will not get the full answer." The NCI came under fire for announcing the findings through a press release, instead of publishing results in a peer-reviewed journal.

## Search for top 50 young biotech groups

[MUNICH] Fifty new biotechnology research groups, each headed by a talented young researcher, are to be set up in Germany next year. Projects will be chosen through a competition called BioFuture which was launched last week by Jürgen Rüttgers, the federal research minister.

An autonomous panel of experts will select the first projects this summer, but young scientists have until January 1999 to apply for the competition. The groups will be small, typically comprising a team leader plus two postdoctoral and two PhD students, and will be hosted by either a university or a research institute. Funding will be limited to five years, and will come from the research ministry's biotechnology budget.

## Scripps and Novartis in genomics link-up

[LONDON] The Swiss-based life sciences company Novartis is to set up one of the world's largest research institutes devoted to functional genomics. The Novartis Institute for Functional Genomics, to be sited adjacent to Scripps Research Institute in La Jolla, California, will receive \$250 million from the company over the next ten years.

Functional genomics involves clarifying the relationship between a particular

genotype and a disease state. Danil Vasella, the president of Novartis, said last week that the company's partnership with Scripps and other academic researchers through the institute "will give us the ability to work simultaneously on a large number of known genetic links, thereby accelerating the pace of our discovery efforts".

## France and Britain ratify nuclear test ban

[LONDON] The United Kingdom and France last week became the first declared nuclear weapons states to ratify the Comprehensive Test Ban Treaty, bringing the total number of ratifications to 13. The treaty has been signed by 149 states since opening for signature in September 1996. It will become effective when ratified by all 44 named states with civilian nuclear capabilities — so far only six have done so.

Three of these named states, India, North Korea and Pakistan, have not even signed the treaty. But its nuclear test verification system is likely to be complete before the treaty enters into force. More than 300 worldwide monitoring stations will be able to detect an air, sea or land-based nuclear explosion with a yield of more than one kiloton.

## Mapping out the path to biodiversity

[LONDON] A group of the world's leading biodiversity scientists, meeting in Mexico last week, have issued recommendations on research needed to implement key articles of the UN biodiversity convention. These include promoting links between scientific institutions involved in cataloguing the world's species, and developing a system of naming species that can be applied in countries with few taxonomists.

Chaired by Ghilleen Prance, director of the Royal Botanic Gardens at Kew in London, the group also recommended research on the interactions between species, and on the impact of human activities on species diversity. The group met under the aegis of Diversitas, a network of scientists, sponsored by six international scientific organizations, that has been given the job of improving links between scientists and the biodiversity convention.

## Mathematics to aid biologists at Princeton

[WASHINGTON] The Institute for Advanced Study at Princeton is to set up a research initiative in theoretical biology, to begin this autumn. The initiative will focus on using mathematics to analyse biological processes, including infectious agents such as HIV, and prions, as well as the use of mathematics to provide insights into evolutionary biology.

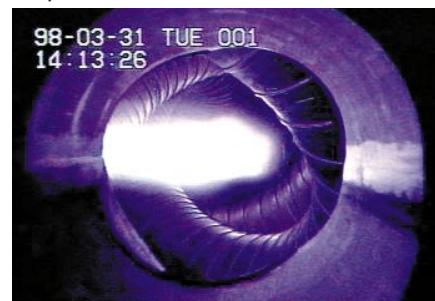
The programme will be headed by Martin Nowak, currently professor of mathematical biology at the University of Oxford. Nowak says that the use of mathematics in biosciences is increasing, and that the Princeton initiative provides a "chance to introduce bright young physicists and mathematicians to a scientific field which is full of open questions and unexplored areas".

## Russian scientists still waiting for pay day

[MOSCOW] Scientists working for the Russian Academy of Sciences last week picketed the offices of Russia's central bank and the Ministry of Finance, demanding payment of their salaries for February and March. The government had transferred the funds to the Moscow National Bank, but the bank was declared insolvent last month, and all accounts have been frozen.

Vladimir Zaitsev, an official from the federal treasury department, has promised to transfer funds to another bank, Rossiysky Credit. He says that the scientists should receive the money before the end of April, and that Tatyana Nesterenko, head of the federal treasury, has agreed to pay the scientists immediately.

## Steady as she goes for Japan's fusion device



[TOKYO] Scientists at Japan's National Institute of Fusion Sciences witnessed the successful first operation of the world's largest stellarator fusion device last week. The Large Helical Device (LHD), which is designed to explore new aspects of steady-state plasma physics, produced the first plasma (pictured) at 1.5 tesla, which is less than half its maximum magnetic energy. The institute aims to bring the reactor up to its maximum capacity over the next six months.

The reactor cost ¥50 billion (US\$373 million) and took eight years to build. Funded by the Ministry of Education, Science, Sports and Culture, it is a key element of Japan's fusion research programme. Support for other components of this programme, such as the JT-60 tokamak and Japan's contribution to the International Thermonuclear Experimental Reactor, comes from the Science and Technology Agency.

# Problems of germline therapy

Sir — You recently published a full-page report of a Californian symposium on germline gene therapy, and a leading article, without a single mention of preimplantation genetic diagnosis<sup>1</sup>.

If a couple are at risk of having a child with a serious genetic disease, it is now possible for them to have their embryos screened at the eight-cell stage, after *in vitro* fertilization, to ensure that only unaffected embryos are transferred to the uterus. Only in the very rare cases where both partners are sufferers from a recessive condition that allows survival to reproductive age, such as cystic fibrosis, will no unaffected embryos be generated. As 10–20 embryos could be produced from a single egg recovery, it would not be difficult also to avoid the birth of carriers if that was desired.

Most couples would surely prefer to avoid the transfer of affected embryos, rather than seeking to tamper with their DNA at such an early stage, with possibly unpredictable consequences. Leroy Hood, chair of molecular biotechnology at the University of Washington, said: "We are using exactly the same kinds of technologies that evolution does". For those who know anything about evolution and its many failures, this is hardly a strong recommendation.

James Watson is reported as saying:

"Scientists should proceed unhindered towards germline engineering". Either he has forgotten that the simpler and safer technique of preimplantation genetic diagnosis, already in clinical use, renders germline gene therapy for genetic diseases virtually pointless, or it is germline engineering for genetic enhancement towards which he wishes to proceed unhindered?

If it is the latter, he should say so. How about it, Jim?

**Anne McLaren**

Wellcome/CRC Institute,  
University of Cambridge,  
Tennis Court Road, Cambridge CB2 1ND, UK  
e-mail: a.mclaren@welc.cam.ac.uk

Sir — One statement in your recent leading article about the consequences of germline gene therapy<sup>1</sup> is unfortunately familiar: "Our first task should be to take a long, hard look at [what] is likely to be involved — both scientifically and ethically". This recalls "Government, religious, civic, and scientific leaders should encourage widespread public discussion of the pros and cons of germ line gene therapy"<sup>2</sup> and, from a Commentary in *Nature*, "timely ethical discussion of this issue, before germline gene therapy in humans is technically feasible, may assist future

policy-makers in their deliberations"<sup>3</sup>.

Indeed, this apparent concern was voiced in one of the first official inquiries into the subject: "The novelty of gene splicing ought not to erect any automatic impediment to its use but rather should provoke thoughtful analysis. Especially close scrutiny is appropriate for any procedures that would create inheritable genetic changes"<sup>4</sup>.

As technical advances make germline gene therapy an even more imminent possibility, one can ask, what has happened in the intervening 16 years? Why has this debate not reached a broader audience? And when will the debate spread beyond the offensive pronouncements of James Watson, who, when once asked if he feared that genetic engineering could be used for 'positive eugenic' ends, replied, "It's not much fun being around dumb people"?

**Jonathan Ewbank**

Centre d'Immunologie de Marseille-Luminy,  
163 Avenue de Luminy, Case 906,  
13288 Marseille Cedex 9, France  
e-mail: ewbank@ciml.univ-mrs.fr

1. *Nature* **392**, 315 & 317 (1998).
2. Institute of Medicine, National Academy of Science. *Human Gene Therapy* (Harvard University Press, 1988).
3. Walters, L. *Nature* **320**, 225–227 (1986).
4. US President's Commission for the Study of Ethical Problems in Medicine and Behavioral Research. *Splicing Life* (US Government Printing Office, Washington DC, November 1982).

## DIY meetings

Sir — Your leading article on postdocs' dissatisfaction with the conference circuit will strike a chord at all levels in the scientific community (*Nature* **392**, 211; 1998). It is not just a problem of competitiveness but rather one of same faces, same stories. The solution is simple — no invited lectures and an intensive programme of 15-minute presentations by the postdocs and postgrads who are doing the experiments. The UK Molecular Microbial Ecology Group is holding its 4th Annual Meeting at the University of Warwick and yet again the scientific programme is a delight.

**Alan J. McCarthy**

School of Biological Sciences,  
University of Liverpool, Liverpool L69 3BX, UK  
e-mail: aj55m@liverpool.ac.uk

## Italian reforms

Sir — Giulio De Leo *et al.* draw attention to the debate in Italy about the reform of research policy (*Nature* **391**, 12; 1998). They also show the results of a

bibliometric analysis of papers in environmental research produced by the universities, CNR, ENEA and other scientific institutions between 1981 and 1996. From these statistics they conclude that the ENEA production is low in numbers and quality.

But ENEA (Italian National Agency for New Technology, Energy and the Environment) is not a typical research institution, because it exists to produce know-how as well as to provide services (under contract), advice and support to the public administration at national, regional and local level. ENEA therefore also produces technical reports and other written and electronic material that is not included in citation indexes.

The field examined by De Leo *et al.* is limited to environmental research. As the title of our agency indicates, environment is only one of ENEA's fields of interest. As a consequence, even within its environment department, research projects and the publications derived from them could easily escape this categorization, because they address areas such as radiobiology, toxicology, geology and numerical modelling, and often yield high-quality

publications that are not always well represented in the journals examined by the authors.

The environment department of ENEA was created in 1994 — or in 1989 under a different name and with different staff. Neither date corresponds to the period covered by De Leo *et al.*

We feel, therefore, that a comparison cannot be made in the oversimplified way followed by the authors. Their approach is flawed by differences in role, history, size, manpower characteristics and fields of publication of the institutions they analyse.

A discussion of the role of ENEA and of the other Italian research bodies is timely and necessary and should consider the distribution of manpower and of financial resources, two items on which discussions should be started and action taken, and De Leo *et al.* are going in the right direction.

**Francesco Mauro**

(Director)  
Environment Department,  
ENEA, CR Casaccia,  
00060 Rome,  
Italy

## No viral disease after xenotransplantation

**Sir**— You refer in your Briefing on xenotransplantation (*Nature* **391**, 320–325; 1998) to a study on seroreactivity to porcine viruses in the ten diabetic patients transplanted with fetal pig islets in Stockholm, Sweden, between 1990 and 1993. The study was presented at the Xenotransplantation Congress in Nantes, France, last September.

Your description of the serological findings is correct: a number of patients tested positive for antibodies to various pig viruses. The sera have subsequently been evaluated for possible cross-reactivity with antibodies to human pathogens, and most reactivities have now been accounted for by such cross-reactivity. There remains a weak reactivity against porcine parvovirus in samples from three patients, and these are being reanalysed at the Centers for Disease Control in Atlanta, Georgia.

The statement that one patient is ill with porcine parvovirus infection is, however, false. None of the patients developed any infectious symptoms related to the xenotransplantation. Nor has any of them developed any unusual symptoms of diseases since exposure to the porcine islets. The alleged transmission of viral disease from pig to man has raised considerable concern. There has, however, been no evidence of such disease in our patients.

**A. Tibell**

**C. G. Groth**

*Department of Transplantation Surgery,  
Huddinge Hospital, S-141 86 Stockholm, Sweden*

## Student cheats

**Sir**— You report that incidents of cheating by students have increased significantly over three decades, amounting to more than two-thirds (73%) of all students polled in US universities without an 'honour code' and about a half (49%) in universities with an 'honour code' and punishments for its violation (*Nature* **390**, 430; 1997).

At the beginning of this century, the record may have been even higher. Abram F. Ioffe (a solid-state physicist) reported in his book *Meetings with Physicists* (Nauka, 1983) that this problem was raised long ago by another prominent scientist, Gilbert N. Lewis (acids-bases theory), with whom he had worked at the University of California at Berkeley in the 1920s. At that time, students forced the university senate to allow them to take written examinations in the absence of instructors.

Lewis proved that 90% of the examination papers were written using

unsanctioned sources. There was also an 'honour' angle. After Lewis's report, the senate ordered the students to add at the end of their work a declaration that they deserved to be trusted in examinations on their honour as American citizens. Lewis proved that this addition did not make a significant difference. (No punishment was incurred with this declaration.)

The senate was furious — with Lewis. It is tempting to suggest that the 90% rate of student cheating found by Lewis is a potential rate achieved under pure experimental conditions (in the absence of instructors and punishments), which can be reduced by instructor presence (to 73%) and further reduced by the application of punishment (to 49%).

Ioffe remarked that, on balance, when he examined his students orally after his course, he had only positive impressions. Thus the inclination to cheat did not seem to interfere significantly with the willingness and ability to learn.

**Alexander Vinogradov**

*Institute of Cytology,  
Russian Academy of Sciences,  
Tikhoretsky Ave 4,  
St Petersburg 194064, Russia  
e-mail: aevin@mail.cytspb.rssi.ru*

## Speculation is premature

**Sir**— Further to your News item about a report being prepared by the National Bioethics Advisory Commission on research involving human biological materials, which prompted some correspondence (*Nature* **391**, 314 & **392**, 14; 1998), I should like to emphasize that the commissioners are still discussing this issue and have neither reached conclusions nor made recommendations.

What do exist are working drafts prepared by staff, papers commissioned by various experts and testimony by the public and professional groups. Speculation about 'likely' recommendations is therefore premature.

The commission expects to receive another staff draft in time for its next meeting on 19–20 May in Cleveland, Ohio. We intend to make this draft available to the public on our Web site (<http://www.bioethics.gov>) before the meeting. The commission is planning to complete and distribute an interim report in the early spring. By distributing these documents widely, the commission hopes to benefit from public and professional comment before issuing a final report.

**Eric M. Meslin**

(Executive Director)

*National Bioethics Advisory Commission,  
6100 Executive Boulevard, Suite 5B01,  
Rockville, Maryland 20892-7508, USA*

## Plight of hawksbill turtles

**Sir**— In reply to my Commentary<sup>1</sup> on the need for open science as the basis for decisions on conservation, Anne Meylan<sup>2</sup> muddles up several different issues about hawksbill turtles, the International Union for the Conservation of Nature (IUCN) and Cuba.

One issue is that it is now close to 18 months since the hawksbill was listed as critically endangered, but documentation has still not been provided. Meylan indicates that because those making the recommendation were specialists there is no need to question it. This appeal to authority, combined with the excuse that her working group are volunteers, will do little to increase confidence that IUCN is achieving the transparency and professionalism that is meant to underlie its categorizations of the status of species. Sanctioning publication of conclusions so far in advance of the supporting analyses is not a standard scientific practice. Meylan claims that data are already available, especially in the Groombridge and Luxmoore report<sup>3</sup>. She fails to mention that those authors suggested that the hawksbill's status could be "indeterminate" rather than "endangered".

A different issue is whether trade in wildlife can be beneficial sometimes. In the case of hawksbills, no fewer will be killed as a result of decisions made at the Convention on International Trade in Endangered Species (CITES) because the Cubans are already taking hawksbills for meat. What the CITES decision means is that the shells from turtles that are already being killed to feed impoverished people will go largely to waste instead of generating support for improved enforcement, monitoring, research, and reduction of incidental catch.

If limited trade is sanctioned under CITES, it actually provides more leverage because it is accompanied by international scrutiny through inspections and reports. Outside CITES, Cuba can do what it wishes with turtles in its waters. The possibility of regulated trade under CITES has already stimulated Cuba to devise an enforcement tool that sets new standards: every piece of hawksbill shell in the Cuban stockpile was videoed and numbered.

**N. Mrosovsky**

*Department of Zoology,  
University of Toronto,  
Toronto, Ontario, M5S 3G5, Canada  
e-mail: nm236@hermes.cam.ac.uk*

1. Mrosovsky, N. *Nature* **389**, 436 (1997).

2. Meylan, A. *Nature* **391**, 117 (1998).

3. Groombridge, B. & Luxmoore, R. *The Green Turtle and Hawksbill (Reptilia: Cheloniidae): World Status, Exploitation and Trade* (Secretariat of CITES, Lausanne, Switzerland, 1989).



# Adapting to change in Latin America

The recent political and economic changes in Latin America provide a challenge to the scientific community. They may be the very stimulus needed to improve the quality of research.

**Fernando C. Reinach**

The countries of Latin America are undergoing rapid and sweeping changes both politically and economically. They now have democratically elected presidents, there has been large-scale privatization, and their economies are open to foreign trade and investment and have stabilized. These changes provide a challenge to the scientific community and the universities. They can either come out of this turmoil reinvigorated as strong players in the economic progress of their country or allow themselves to put aside as inefficient machinery incapable of responding to new demands. After many years of stagnation and resistance to previous governments, it is important that the universities and the scientific community as a whole are prepared to respond to these challenges.

## Historical problems

During the 1970s and 1980s, most of the countries in the region were under military rule. Their economies were largely isolated because of high import restrictions to protect local producers who were not competitive with foreign competition. Industrial production was based on outdated technology imported after it had become obsolete in developed countries. In Brazil, for example, a personal computer cost seven times the price in the United States for a model that was out of production there, and most new cars had engines developed in the mid-1950s. State-owned companies financed with borrowed money were guaranteed monopoly by law; these included most of the oil and steel industries, telecommunications and power generation.

In the late 1980s this economic model collapsed throughout the region. Foreign investment and inflation grew out of control, the technology increased, and the state-owned companies were shown to be incapable of fulfilling their roles. Anyone needing to buy a tele-

phone in Brazil could buy it on the black market for around \$3,000 or could wait in line for up to three years. Inflation rates of 1,000% a year were not uncommon in Brazil, Argentina and Chile.

There was no real economic need to invest in science, but the rhetoric of 'science-based development' kept science alive. This rhetoric fulfilled the nationalistic objectives of governments and was used by local industries to keep the markets closed. It was also useful to scientists: it kept the money flowing and allowed relatively large sci-

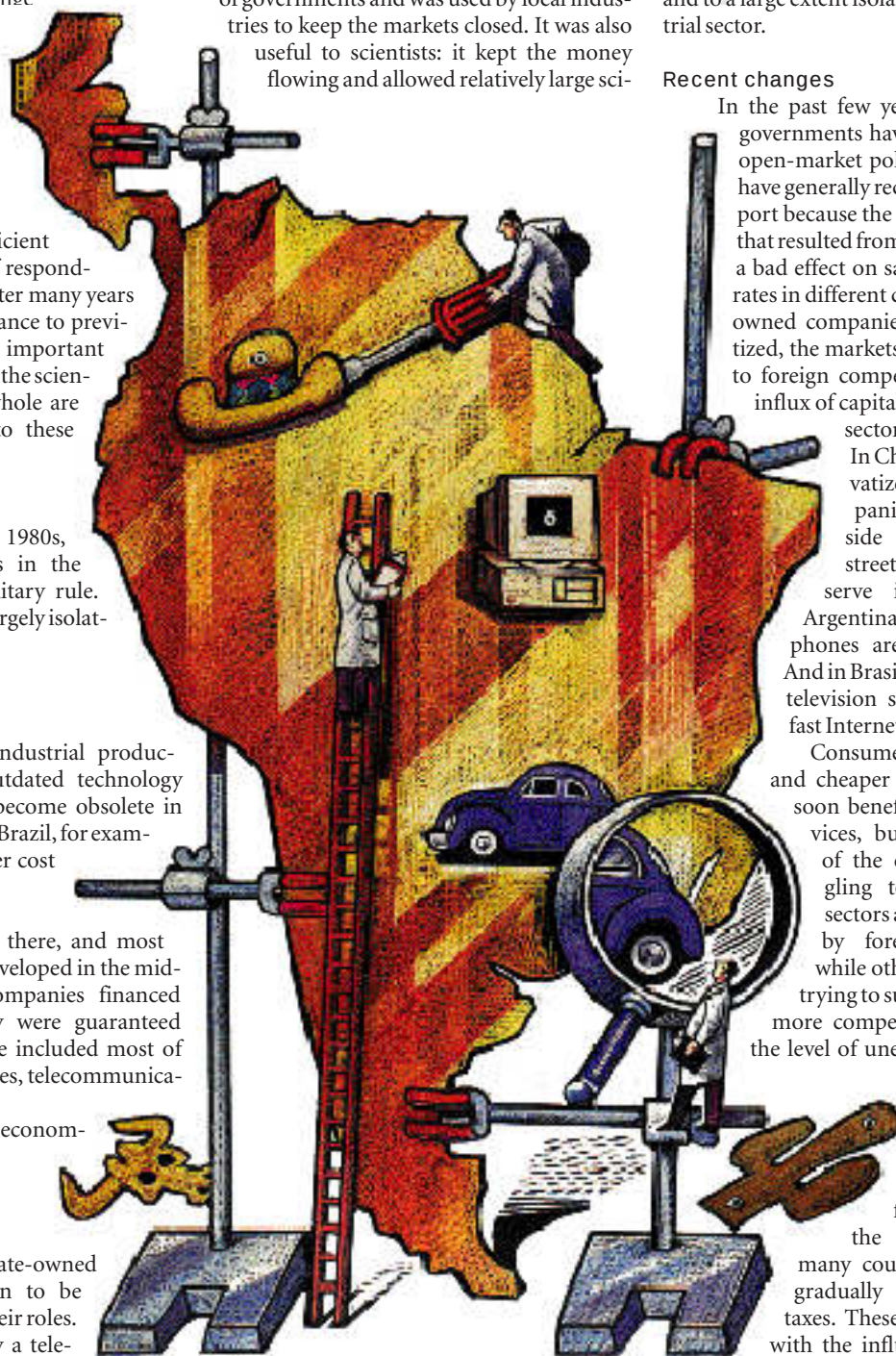
entific communities to grow in many Latin American countries. Graduate schools developed in the main universities, and many students were sent abroad for PhDs or postdoctoral work. Basic science survived and grew with the limited funds available. As a result, Brazil, Argentina, Chile and Mexico all have sizeable scientific communities, based mainly in the state-owned universities and to a large extent isolated from the industrial sector.

## Recent changes

In the past few years, newly elected governments have implemented an open-market policy. These changes have generally received popular support because the high inflation rates that resulted from the old system had a bad effect on salaries. At different rates in different countries, the state-owned companies are being privatized, the markets have been opened to foreign competition and a large influx of capital is reshaping many sectors of the economy. In Chile, competing privatized telephone companies have rival booths side by side on the streets. Cellular phones serve isolated areas in Argentina where land-line phones are still unavailable. And in Brasilia the regular cable television system comes with fast Internet connection.

Consumers can enjoy better and cheaper products and will soon benefit from better services, but a large fraction of the economy is struggling to survive. Whole sectors are being wiped out by foreign competition, while others are desperately trying to survive by becoming more competitive. As a result, the level of unemployment is at a record high in Argentina and is growing steadily in Brazil.

In an attempt to force competition, the governments of many countries have begun gradually to reduce import taxes. These reductions, along with the influx of capital from



PETER BREWSTER

foreign companies constructing modern production facilities, will force the outdated production sector to modernize.

## The challenge ahead

I believe that economic pressure will change the relationship between science and society. The question is whether the scientific community is willing and prepared to make the best of the changes.

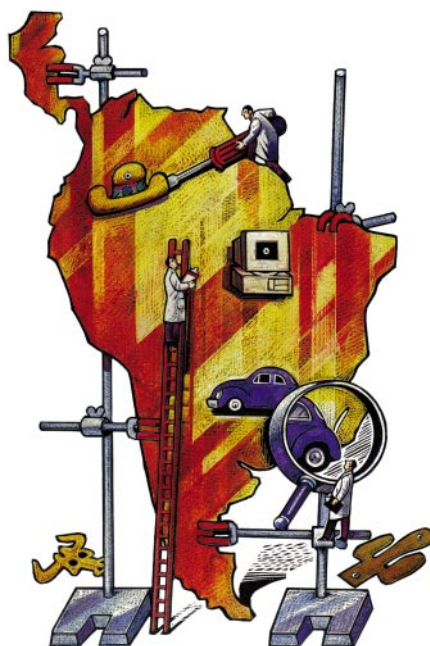
It is clear that much of the demand for new technology will be supplied by importation from abroad, as exemplified by the recent agreement between the Massachusetts Institute of Technology and the federal granting agency Financiadora de Estudos e Projetos FINEP in Brazil. But it is also clear that part of this demand will converge on the universities and the scientific community, if for no better reason than that the industrial sector expects some return on its taxes — in Latin America most universities are supported fully by the government. In São Paulo, Brazil, state universities receive around 10% of the state taxes; sooner or later they will have to have something to show for it.

This demand will force us to confront economic reality. Can universities continue in complete isolation from the industrial sector, concentrating all their research effort on basic science teaching? Or should they instead spend a fraction of their effort in the generation and transfer of new technology to the industrial sector? Although most scientists agree that they should make this change, few realize that it will force universities down a road that will drastically change their nature.

To move in this direction, universities will have to learn a whole new set of skills. Institutions where consultancy work is currently forbidden will have to learn how to regulate it while preserving academic quality and basic science. An environment that has recently learned to value publication as a measure of productivity will now have to learn to evaluate technological projects where publication is sometimes irrelevant. These issues, which are part of the daily lives of scientists in Europe and the United States, are still not being discussed in Latin America.

## Effects of change

If the university is not to be transformed into an institution devoted only to applied research, and strong basic science is to be maintained alongside research efforts directed at providing technology for the industrial sector, then the perceived value of science must increase. Such changes are already apparent in many countries. In Chile, for example, university faculties are expected to generate part of their income from consultancy work. Financing bodies, such as FAPESP and PADCT in Brazil, have programmes that support projects based on direct collaborations between industry and



The scientific establishment in Latin America is like a species confronted with a large change in environment: if flexible enough to adapt it will evolve, otherwise it may face extinction.

universities. In Argentina, many small biotechnology companies have been set up by scientists, a move given further impetus by a period of very low wages.

But if the scientific community does not respond well to the changes, there may be increased resistance towards these changes from the universities or, even worse, governments may pull out from the universities before the changes have time to happen. Signs of this trend are already being seen in Brazil: the government, frustrated with its inability to rationalize spending in the federal universities and without the political power to impose selective cuts on the worst units, is slowly squeezing the budget across the whole sector. These signs are particularly worrying, because in Latin America, state-owned institutions are rarely extinguished but rather left to die through terminal neglect and lack of funding.

Economic pressure will also change the way funds are allocated to universities and scientific research; this is presently widely recognized as being wasteful. The present system pays salaries to non-productive staff and results in poor peer review of scientific projects and a bloated and inefficient admin-

istrative infrastructure. It has been argued that the percentage of the gross national product spent on science in Latin America, although lower than that spent in developed countries, is sufficient to support the growth of science if it is targeted at the relatively few productive scientists. Despite this, universities have always been compared with state-owned companies that were equally badly managed. The effects of privatization and the consequent increase in productivity will remove this comparison and raise public awareness, and hopefully create a strong demand for increased accountability and clear indicators of efficiency.

If successful, this process may result in a remodelled administrative structure with an end to early tenure decisions (some universities in Brazil grant tenure with full salary before a PhD is obtained), the firing of unproductive faculty and staff, the trimming of administrative costs, and increased financial independence of universities through endowments and other forms of financing.

But these are difficult changes. Most scientists in Latin America are poorly paid public servants who have a job for life, a salary paid by the government, and guaranteed early retirement on full salary. At the University of São Paulo, about half of the salary bill goes to retired faculty and staff. These privileges, unknown in developed countries, are considered important by staff and have historically helped to guarantee a degree of freedom of speech in the universities. Are faculty members prepared to trust a democratic government and exchange these privileges for a higher salary?

Again, if the universities resist the inevitable change they may fare badly. If the universities are deemed to provide poor value for money, and privatization continues to reap rewards, society might decide it can live without public universities. Governments may then be tempted to privatize the whole university system or replace it with private, for-profit teaching universities.

I believe that science in Latin America is at a difficult crossroads. As an optimist, I see the situation as a unique opportunity. The economic changes offer the right combination of pressure and incentives to allow dramatic improvement in the quality of the science produced in Latin America and in its possible effect on the development of the region. This chance must not be wasted. The scientific establishment in Latin America is like a species confronted with a large change in environment: if flexible enough to adapt it will evolve, otherwise it may face extinction. □

*Fernando C. Reinach is in the Department of Biochemistry and the Institute of Chemistry, University of São Paulo, Caixa Postal 26.077, 00559-970 São Paulo, SP, Brazil. e-mail: fdcreinach@iq.usp.br*



# In victu veritas

Harold McGee

A survey of spice use around the world concludes that spices serve the adaptive purpose of reducing food-borne disease. It highlights, however, the need for further research *in victu* — in food itself — rather than *in vitro*.

Like all animals, human beings eat to live; and food preparation is a cultural behaviour that presumably contributes to our fitness by making plant and animal tissues more nourishing. But food preparation in many cultures has become far more elaborate than simple survival would seem to require. Why do we invest so much effort in adorning and transforming our cereals, tubers, meats and milks? In particular, why do humans bother to flavour their foods with nutritionally insignificant quantities of herbs and spices? And why does tropical heat seem to foster especially pungent cuisines?

In an extensive literature survey and correlational study, published in *The Quarterly Review of Biology*<sup>1</sup>, Billing and Sherman affirm the relationship between climate and spiciness. They attribute it to what they consider the primary adaptive value of seasoning — that antimicrobial compounds concentrated in spices reduce the incidence of food-borne disease.

Drawing on 93 cookbooks covering 36 countries, Billing and Sherman analysed 4,578 meat, poultry and seafood recipes for the number and kinds of spices included. (They use 'spice' to signify all plant flavourings, whether Far Eastern natives such as pepper and cloves, central Asian garlic and onions, Mediterranean herbs such as thyme and oregano, or central American chillies. Quantities were not considered.) They found a strong correlation between the mean annual temperature of a given country — an index of the rate at which foods will spoil there — and the mean number of spices added to its flesh dishes, which ranged from two in Norway and three in Ireland to ten in India. The authors also surveyed the literature on the antibacterial activity of plant-derived flavourings (which is due largely to phenolic, terpenoid and sulphur compounds), and found a correlation between a country's climate and the proportion of its recipes calling for onion, garlic and chillies, and 12 other 'highly inhibitory' spices. Both onion and pepper (a weak inhibitor) appeared in nearly two-thirds of recipes worldwide, with garlic and then chillies, lemon and lime juice, and parsley trailing well behind.

Billing and Sherman considered the 'null hypothesis' that people simply embellish their food with locally available flavourings,

but found no significant correlation between the production of a spice in a given country and its consumption in that country. They also reject the adaptiveness of using spices to disguise spoilage, which they point out would increase the likelihood of food poisoning. They conclude that although the proximate reason for spice use is to make food more palatable, the 'ultimate reason is most likely that spices help cleanse food of pathogens and thereby contribute to the health, longevity, and reproductive success of people who find their flavors enjoyable'.

Food historians and other writers have often casually remarked that spices help preserve foods, so a systematic approach to the subject is most welcome. But correlational studies can only be as reliable as their data sets, and the data chosen in this study are not well suited to shed new light on spicy heat.

The large recipe database turns out to be a narrow source of evidence for the 'ultimate' purpose of an ancient habit across all cultures and historical periods. No cookbook consulted ante-dates 1945, by which time any biologically determined patterns of spice use have long been obscured by migrations of peoples and plants, technical advances in agriculture, transport and food handling, and other social and economic factors. Formerly rare tropical spices are now readily available in northern supermarkets alongside local herbs; conversely, black pepper is little-used in its native south India because it is a lucrative export crop, and chillies are a cheap alternative. Given these and similar considerations — for example, that India and its warm neighbours no doubt grow and consume a broader range of all food plants than does northern Europe, not just spices —

the null hypothesis survives pretty well unscathed.

A second limitation of the recipe database is that the authors consulted only English-language cookbooks. The ingredient lists in African and Asian books appear to have been modified for Western readers and their larders (no indigenous African spices make the list, nor do Japanese wasabi or shiso, or Chinese star anise or Szechwan pepper, which is neither *Piper* nor *Capsicum*). Together with the exclusive emphasis on meat and fish dishes, this Western orientation may also give undue weight to recipes of the affluent few who are best able and most likely to use ancillary ingredients. For example, an Indian study<sup>2</sup> found that rural labourers near Hyderabad tend to spice their predominantly cereal regimen more heavily than the urban middle-class did its broader diet, but with a more limited palette of chillies, tamarind and turmeric. In Billing and Sherman's tabulation of spice use, this sample of the average Indian diet would rank closer to Ireland than to the India of the cookbooks.

The most serious weakness in Billing and Sherman's case for their hygienic hypothesis is the way in which they represent the antibacterial activities of spices. Studies in this area have used a great variety of experimental conditions.

The authors there-



ANTHONY BLAKE PHOTO LIBRARY/MERHURST



fore chose simply to tabulate the percentage of all bacteria tested towards which a given spice or its extract has consistently shown at least some inhibitory activity, regardless of dose. Four spices are said to be inhibitory towards all bacteria tested — garlic, onions, allspice and oregano — and a total of 15 inhibited 75% of bacteria, or more.

This simple ranking of antimicrobial potencies disguises large differences in the extent to which different spices have been tested. An appendix reveals that in the most inhibitory group, garlic was tested against 29 bacteria, but allspice against only five, and onion against four. Chillies were successful against just four of five bacteria. A review of chemical food preservation concludes that of the 15 spices that Billing and Sherman classify as 'highly inhibitory' and find significantly associated with warm climates, seven, including chillies and allspice, have shown "limited or no activity" towards food spoilage microbes<sup>3</sup>. Even for the apparently inhibitory spices (cinnamon and cloves, garlic, and the Mediterranean herbs), the problem remains that most evaluations have been done *in vitro* rather than *in victu*. The few spices to have been tested in foods show far less activity than they do in microbiological media, probably because food particles sequester the inhibitory compounds<sup>3</sup>. The active sulphur compounds in garlic and onions are formed in disrupted tissue by enzymatic activity, which ordinary cooking procedures limit (but which the Indian practice of grating or grinding probably maximizes). And little is known about the stability of the inhibitory compounds to cooking temperatures.

In short, we don't yet have enough information to say whether spices have significant antimicrobial effects in ordinary cooking. On balance, it looks as though a simpler and more reliable defence against foodborne disease is to heat food well and often, as Indians have long done with highly perishable and unspiced liquid milk.

If Billing and Sherman are unable to make a convincing case for the hygienic hypothesis, their labours are to be valued for reviving an excellent question and demonstrating the need for *in victu* research initiatives. After testing culinarily appropriate spice doses in a model food, one would want to contaminate genuine, pre-refrigeration versions of, say, chicken in a sauce — Chinese red-cooked, Indian curried, Italian *cacciatore*, French *coq au vin*, American pot pie, Mexican *mole* — and follow pathogen growth in each, which will surely be affected by ingredients such as soy sauce, vinegar and wine. Given the worldwide popularity and yet limited antibacterial activity of black pepper, Billing and Sherman's speculative rationale — that pepper acts as a 'bioavailability enhancer' to sensitize bacteria to other spices — needs testing, as does the possibility that

common spice combinations bring together complementary antimicrobials.

The strong biological slant of Billing and Sherman's hygienic hypothesis will also help sharpen the thinking of anyone who is intrigued by the complexities of human spice use and its persistence. The past few decades have brought Paul Rozin's rich behavioural analysis of flavourings and the challenges of omnivory<sup>4</sup>, fascinating studies of the ceremonial and mythopoetic<sup>5</sup>, social<sup>6</sup> and economic<sup>7</sup> aspects of European spice habits — and a continuing boom in per capita spice consumption in the United States. Is the last a manifestation of new selection pressures exerted by *Escherichia coli*, salmonella and campylobacter? Of an immigrant population reverting to type? The chemosensory

circuits of my salsa- and vindaloo-loving family are a legacy of unspiced northern Europe, and they suggest not. □

Harold McGee is at 14 Chemin Las Crozes, 31450 Pompertuzat, France. He is author of *On Food and Cooking: The Science and Lore of the Kitchen*, and *The Curious Cook* (Scribner's/Allen & Unwin).

e-mail: [hjmcgee@alumni.caltech.edu](mailto:hjmcgee@alumni.caltech.edu)

1. Billing, J. & Sherman, P. W. Q. *Rev. Biol.* **73**, 1–38 (1998).
2. Thimmayamma, B. V. S., Rau, P. & Radhaiah, G. *Ind. J. Nutr. Dietet.* **20**, 153–162 (1983).
3. Davidson, P. M. in *Food Microbiology* (eds Doyle, M. P., Beuchat, L. R. & Montville, T. J.) 520–556 (Am. Soc. Microbiol., Washington DC, 1997).
4. Rozin, P. in *Flavor: Its Chemical, Behavioral, and Commercial Aspects* (ed. Apt, C. M.) 101–127 (Westview, Boulder, CO, 1978).
5. Detienne, M. *The Gardens of Adonis*, transl. Lloyd, J. (1972; reprint Princeton Univ. Press, 1994).
6. Schivelbusch, W. *Tastes of Paradise* (Pantheon, New York, 1992).
7. Laurioux, B. *Food Foodways* **1**, 43–76 (1985).

## Mantle mineralogy

# Olivine emerges from isolation

Craig R. Bina

Near a depth of 410 km, the speeds at which seismic waves travel in Earth's mantle rise sharply, an observation long attributed to high-pressure transformation of the mineral olivine to wadsleyite. However, seismological studies of this '410-km discontinuity' reveal wave-speed increases that are smaller and more abrupt than expected from extrapolation of the laboratory behaviour of these minerals to mantle conditions. On page 702 of this issue<sup>1</sup>, Irifune and Isshiki show that the hitherto largely neglected exchange of

magnesium and iron between olivine and other mantle minerals is an important factor in resolving these apparent discrepancies.

Such discrepancies are significant because they constrain the extent to which mantle discontinuities arise from phase transitions in minerals, such as those in olivine, rather than changes in bulk chemical composition. The mineralogy and chemistry of Earth's mantle, especially that part known as the 'transition zone' between the seismic discontinuities at depths of 410 and 660 km, govern the dynamics of convective

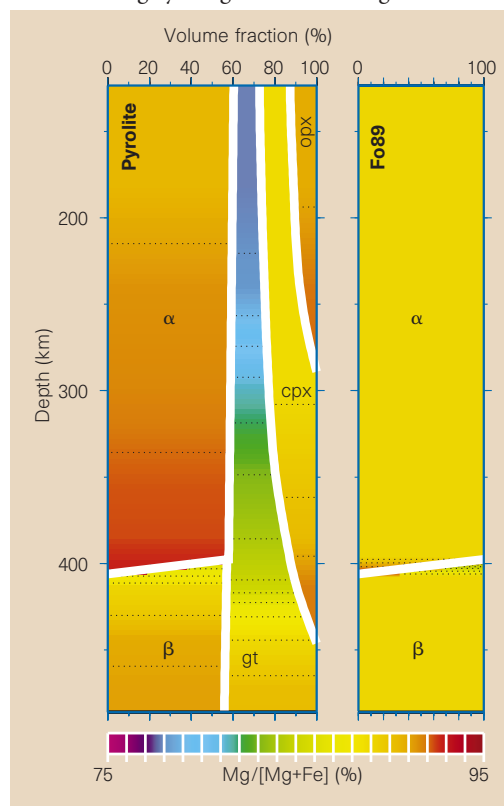


Figure 1 Equilibrium mineral proportions and compositions as functions of mantle depth, for  $(\text{Mg}_{0.89}\text{Fe}_{0.11})_2\text{SiO}_4$  olivine in isolation (Fo89) and for a model (pyrolite) mantle composition. Mineral proportions are as volume fraction, shown by bold white lines; compositions are in terms of  $\text{Mg}/[\text{Mg}+\text{Fe}]$ , and are shown by colours and by dotted contours at 1% intervals. With increasing depth, olivine ( $\alpha$ ) transforms to wadsleyite ( $\beta$ ) near 410 km, and orthopyroxene (opx) gradually dissolves into clinopyroxene (cpx) which in turn dissolves into garnet (gt). In the pyrolite mantle composition, all phases grow more Mg-rich with increasing depth. (Figure based on data from Irifune and Isshiki<sup>1</sup>.)

flow of heat and matter in the interior<sup>2</sup>. This flow, in turn, drives the tectonic evolution of the surface, including the occurrence of volcanism and seismicity.

Olivine, a solid solution of forsterite ( $\text{Mg}_2\text{SiO}_4$ ) and fayalite ( $\text{Fe}_2\text{SiO}_4$ ), has the approximate composition  $(\text{Mg}_{0.89}\text{Fe}_{0.11})_2\text{SiO}_4$  (called forsterite-89 or Fo89) in samples from the shallow mantle. It transforms from the olivine ( $\alpha$ ) phase to wadsleyite ( $\beta$ ), changing again to ringwoodite ( $\gamma$ ) at greater depths and eventually breaking down to a mixture of silicate perovskite and magnesiowüstite. In standard compositional models of the upper mantle such as the pyrolite<sup>3</sup> model, in which a mix of different minerals is taken to account for the rocks and melts erupted at the surface from the interior, olivine accounts for only about 60% of the mantle by volume. Other minerals concerned include orthopyroxene, clinopyroxene and garnet, which gradually dissolve each into the next with increasing depth. Despite a pioneering study nearly 20 years ago, in which Akaogi and Akimoto<sup>4</sup> pointed out that knowledge of element partitioning amongst these minerals is 'indispensable' for understanding the mantle, to this day the olivine polymorphs are treated in isolation from pyroxenes and garnets in most models of the mantle.

Irifune and Isshiki<sup>1</sup> have shown such artificial separation to be misleading by experimentally demonstrating Mg-Fe exchange between  $\alpha$ ,  $\beta$ , garnet and clinopyroxene. We have known<sup>5</sup> for some time that the proportions of these minerals vary with depth, as shown by the bold white lines in Fig. 1, but Irifune and Isshiki have now measured the accompanying variations in the compositions of coexisting minerals, shown by the colours. In isolated Fo89 olivine, mineral compositions must remain constant, except in the narrow zone of  $\alpha$ - $\beta$  coexistence. In pyrolite, however, all phases grow increasingly Mg-rich with depth. Given that bulk composition is fixed, this entails an increasing abundance of the most Fe-rich phase, garnet. Indeed, gradual dissolution of the less Fe-rich pyroxenes into the garnet both raises the proportion of garnet and dilutes the garnet to lower Fe levels. However, given constant Mg-Fe partitioning ratios (incidentally confirmed by Irifune and Isshiki), such dilution drives Fe-depletion (that is, Mg-enrichment) of  $\alpha$  to maintain equilibrium levels of Fe in garnet. As a result, olivine above the  $\alpha$ - $\beta$  transition in pyrolite becomes enriched in Mg relative to that in Fo89.

Such Mg-enrichment of  $\alpha$  shifts the  $\alpha$ - $\beta$

transition to higher pressures (as can be seen in Irifune and Isshiki's Fig. 3 on page 704), so the onset of the transition in pyrolite is postponed to greater depths relative to that in Fo89. On the other side of the phase change,  $\beta$  does not partition Fe into garnet as strongly as does  $\alpha$ , so  $\beta$  is not initially Mg-enriched, and the completion of the  $\alpha$ - $\beta$  transition is not postponed relative to Fo89. The net effect of postponing onset but not completion is to concentrate the transition into a narrower range of depths, decreasing the depth-extent of the wave-speed increase by some 6 km relative to that expected for Fo89 (>10 km), in better agreement with seismic<sup>6</sup> estimates (<5 km).

Furthermore, mineral wave-speeds themselves also vary with composition. Wave-speed increases across the  $\alpha$ - $\beta$  transition in isolated Fo89 can be measured in the laboratory, extrapolated to mantle conditions, and scaled for a mantle model which is about 60% olivine by volume. Although subject to remaining uncertainties<sup>7</sup> in the pressure- and temperature-dependence of mineral wave-speeds, some such analyses<sup>8</sup> yield increases whose magnitudes are rather larger than those inferred for the 410-km seismic discontinuity. In pyrolite, on the other hand, the Mg-

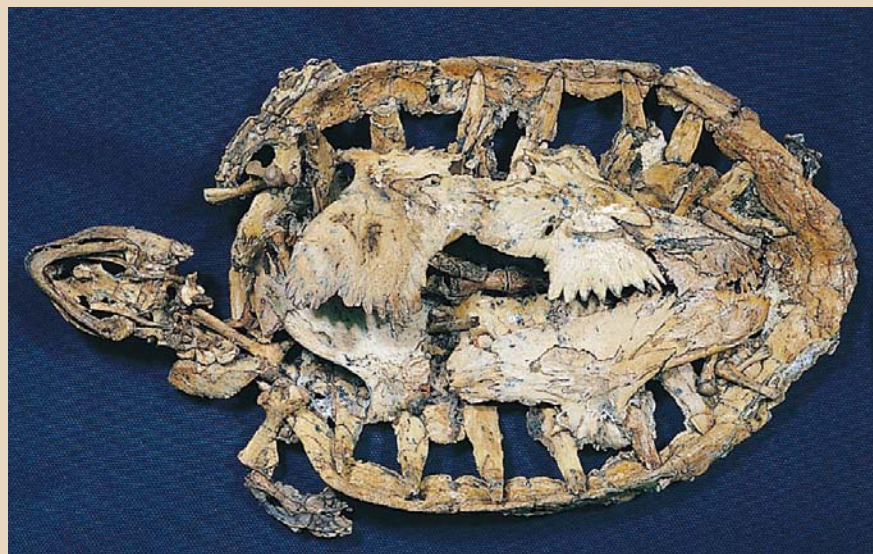
## Palaeontology

### The eyes have it

Life on land is so challenging that many land vertebrates have taken the plunge and gone back to the sea. From porpoises to placodonts, pliosaurus to pinnipeds, the examples are legion. But life in the sea poses a unique problem for a returning land vertebrate — salt.

A marine vertebrate with a terrestrial ancestry has body fluids that are much less salty than sea water, so an unprotected vertebrate immersed in the sea runs the risk of dehydration. Water can be replaced by drinking sea water. This raises the saltiness of the body with respect to the sea, so there can be no net gain of water unless the salts are excreted in a solution at least as concentrated as that of sea water. The kidneys of marine reptiles, such as turtles, do not have this concentrating power: without other methods of voiding excess salt, a marine reptile cannot eat salty food or drink sea water without becoming dehydrated in the midst of plenty.

Marine life is simply unliveable unless an animal solves the salt problem. This 200-millimetre-long, 110-million-year-old fossil marine turtle, described by Ren Hiryama on page 705 of this issue (*Nature* 392, 705–708; 1998), has taken this lesson on board. The creature, which comes from



the Lower Cretaceous Santana Formation of north-east Brazil, extends the fossil record of marine turtles back ten million years. The turtle is primitive in the sense that the bones in its wrists, ankles and digits have not become consolidated into rigid paddles. In other words, its feet would have looked more like those of freshwater terrapins than fully seagoing turtles.

Its salt-excreting arrangements were, in

contrast, far less haphazard. Marine turtles have lachrymal glands (each one larger, in some cases, than the brain) modified to excrete a concentrated salt solution. The skull of the Santana turtle shows evidence that it had enormous salt glands around the eyes. In which case, the evolution of salt-excreting glands preceded that of rigid paddles. These salt glands were the turtles' passport back to the sea.

Henry Gee



enriched  $\alpha$  is seismically faster than the  $\alpha$  in Fo89, but its non-enriched  $\beta$  exhibits the same speeds as  $\beta$  in Fo89. Thus the Mg–Fe partitioning which leads to a narrower transition yields one which is also smaller in magnitude (by about 10% according to Irifune and Isshiki's estimate), in better agreement with seismic estimates.

So Irifune and Isshiki have demonstrated that the olivine polymorphs exchange Fe and Mg with their pyroxene and garnet companions under mantle conditions rather than residing in isolation, confirming Akaogi and Akimoto's early indications that Mg–Fe partitioning relative to garnet differs significantly among the olivine polymorphs. Although other factors, such as mineral hydration or nonlinear variations in wave speed, may also affect the seismic signature of phase changes, Irifune and Isshiki have helped to reconcile the properties of the  $\alpha$ – $\beta$  transition in a pyrolite model mantle with observations of the 410-km seismic discontinuity.

Although their study pertains largely to the top of the transition zone, other work<sup>9</sup> suggests that similar phenomena — Al–Fe exchange involving  $\gamma$ , garnet, silicate perovskite and magnesio-wüstite — may in

part determine the properties of the 660-km seismic discontinuity at the bottom of the transition zone. Interestingly enough, Akaogi and Akimoto's early study also suggested that Mg–Fe partitioning between garnet and  $\gamma$  should be reversed relative to  $\beta$  in the middle of the transition zone. Future work may or may not show that this behaviour does indeed occur. If it does, it could affect the properties of the broader  $\beta$ – $\gamma$  transition, too, with implications for our understanding of the enigmatic seismic reflectors<sup>10</sup> that occur at around 520 km depth. □

Craig R. Bina is in the Department of Geological Sciences, Northwestern University, Evanston, Illinois 60208-2150, USA.

e-mail: craig@earth.nwu.edu

1. Irifune, T. & Isshiki, M. *Nature* **392**, 702–705 (1998).
2. Christensen, U. *Annu. Rev. Earth Planet. Sci.* **23**, 65–87 (1995).
3. Ringwood, A. E. *Composition and Petrology of the Earth's Mantle* (McGraw-Hill, New York, 1975).
4. Akaogi, M. & Akimoto, S. *Phys. Earth Planet. Inter.* **19**, 31–51 (1979).
5. Ringwood, A. E. *Spec. Publ. Geol. Soc. Australia* **14**, 457–485 (1989).
6. Neele, F. *Geophys. Res. Lett.* **23**, 419–422 (1996).
7. Li, B., Gwanmesia, G. D. & Liebermann, R. C. *Geophys. Res. Lett.* **23**, 2259–2262 (1996).
8. Duffy, T. S., Zha, C. S., Downs, R. T., Mao, H. K. & Hemley, R. J. *Nature* **378**, 170–173 (1995).
9. Wood, B. J. & Rubie, D. C. *Science* **273**, 1522–1524 (1996).
10. Shearer, P. M. *J. Geophys. Res.* **101**, 3053–3066 (1996).

## Crop genetics

# Reducing transgene escape routes

Alan J. Gray and Alan F. Raybould

Genetically modified crops and the food derived from them have had a bad press, but a rare piece of good news is provided by Daniell *et al.*<sup>1</sup> in this month's *Nature Biotechnology*. They report a development that has profound implications for the risk assessment of genetically modified crops. Most crops are modified by inserting genes into the nucleus, and the genes can therefore spread to other crops or wild relatives by movement of pollen. By engineering tolerance to the herbicide glyphosate into the tobacco chloroplast genome, however, the authors have not only obtained high levels of transgene expression, but, because chloroplasts are inherited maternally in many species, they have also prevented transmission of the gene by pollen — closing a potential escape route for transgenes into the environment.

Glyphosate is the most widely used herbicide in the world. It interferes with 5-enolpyruvyl shikimate-3-phosphate synthase (EPSPS), an enzyme that is encoded by a nuclear gene and catalyses a step in the biosynthesis of aromatic amino acids in the chloroplasts. Conventional strategies for producing glyphosate-tolerant plants are to insert, into the nucleus, an EPSPS gene from a plant or a glyphosate-tolerant bacterium (the bacterial gene is modified so that the enzyme

is correctly targeted to the chloroplasts), or a gene that inactivates the herbicide.

But Daniell *et al.* have now achieved maternally inherited glyphosate tolerance in tobacco by particle bombardment of leaves with a vector containing the EPSPS gene of

petunia. The gene was modified with border sequences that allow integration into the chloroplast genome by homologous recombination and, using the polymerase chain reaction and Southern blot analysis, the authors confirmed that the gene had inserted into chloroplast DNA of regenerated plants. All of the chloroplasts were transgenic, and the very high copy number of the  $\epsilon$  enc. (up to 10,000) resulted in a tenfold increase in glyphosate tolerance compared with transgenic plants that carry the same gene in the nuclear genome<sup>2</sup>.

Genes that confer tolerance to herbicides have been the target of genetic engineers for more than a decade. Up to January this year, there were nearly 2,300 experimental releases of herbicide-tolerant crops in OECD countries (about 35% of all genetically modified crop releases up to then)<sup>3</sup>. As well as its agronomic objective, herbicide tolerance has been widely used as a selectable marker, enabling the easy detection of successful transformations.

Herbicide-tolerant crops such as the glyphosate-tolerant soybean *Glycine max* were among the first genetically modified species to be grown commercially in North America. Such crops are also near to market in Europe — for example, both glyphosate- and glufosinate ammonium-tolerant oilseed rape, *Brassica napus*, are undergoing variety trials in the United Kingdom. Tolerance to glyphosate has featured in almost one-third of all herbicide-tolerance field trials and in 20 different crops, ranging from lettuce and tomato to eucalyptus, chestnut and poplar, although the major crop targets have been maize, soybean, oilseed rape, sugar beet and cotton.

The environmental and agricultural implications of the large-scale cultivation



Figure 1 Off the beaten track — sea beet on a sea wall in Somerset (UK), flowering at the same time as nearby sugar beet. Genes from sugar beet are known to transfer to wild, weedy beets in areas where sugar beet seed is produced<sup>12</sup>. The same process can occur when sugar beet plants flower before harvest in coastal regions where sea beet occurs.





Figure 2 Bee on *Brassica* flowers. The work of Daniell *et al.*<sup>1</sup> should prevent the transmission of transgenes through pollen.

of herbicide-resistant crops have been hotly debated<sup>4</sup>. Although disagreeing about the possible effects on the use of herbicides, environmentalists and industry both recognize the potential problems created by gene flow to wild relatives and, specifically, the creation of herbicide-tolerant weeds. These problems are addressed by regulators such as the UK Advisory Committee on Releases to the Environment in their risk assessment of genetically modified crops.

Hybridization with wild relatives may cause problems in a number of ways: by the creation of 'superweeds'<sup>5</sup>; by genetic erosion (particularly in centres of diversity); and — a somewhat purist concept — through genetic pollution of natural gene pools. Crops from which gene flow to wild relatives is known to have occurred include quinoa, squash, carrot, maize, sorghum, sunflower, strawberries and sugar beet<sup>5,6</sup> (Fig. 1). Particular problems with gene flow are anticipated in outbreeding crops such as forage grasses, which have extensive dispersal of pollen and many sexually compatible wild relatives. In the case of glyphosate tolerance, preventing gene flow to weedy relatives is especially worthwhile because tolerance in weeds does not seem to have evolved, despite extensive use of the herbicide for over 20 years<sup>7</sup>. By contrast, resistance to more than 15 herbicide groups has been observed in well over 100 species of plant worldwide<sup>8</sup>.

As well as preventing escape to wild relatives, maternal inheritance of transgenes would prevent (or reduce) gene 'stacking' by cross-contamination of crops — an advantage that is not mentioned by Daniell *et al.*<sup>1</sup>. For example, oilseed rape plants tolerant to both glyphosate and glufosinate ammonium could appear where genetically modified crops are grown adjacently. The ability to prevent the contamination of food crops with genes involved in the production of industrial and pharmaceutical products is also of great potential benefit.

Of course, maternal inheritance will not completely prevent the escape of transgenes. Paternal or biparental inheritance of chloro-

plast DNA is common in gymnosperms (plants in which the seed is not protected by an ovary), and this also occurs in several genera of angiosperm (flowering plants)<sup>9</sup>. Furthermore, like conventional crops, transgenic plants will be able to form volunteer populations — plants that arise in fields after harvest. Such populations can cause economic losses because of direct competition with the intended crops for resources such as light and water, and because they can act as 'green bridges' (hosts for pests and pathogens while crop plants are absent). Nor will maternal inheritance stop transgenic crops from establishing 'feral' populations. In the United Kingdom, for example, oilseed rape is notorious for producing such populations on roadsides because of seed spillage during transport to processing factories<sup>10</sup>. Moreover, the seed can persist in the soil for several years and germinate after the ground is disturbed<sup>11</sup>.

Although one of the escape routes for genes from genetically modified crops has been cut off with the work of Daniell *et al.*, not all of the exits are blocked (Fig. 2).

## Mathematics

# How to melt if you must

Ivar Ekeland

Physics has been a rich source of inspiration for mathematics. Even crude models of physical phenomena can lead to systems of partial differential equations (PDEs) that challenge existing mathematical techniques. This is because, until the 1970s, most of the theory of PDEs was developed in a linear framework. Now, as nonlinear models become increasingly popular, mathematicians are striving to keep up by developing techniques appropriate to nonlinear situations. A typical instance is the recent proof by Nassif Ghoussoub and Changfeng Gui<sup>1</sup> of a conjecture by de Giorgi

Nevertheless, the option to prevent transgene movement via pollen offers scope for simplifying the risk assessment of transgenic crops with strict maternal inheritance of chloroplasts — particularly those species in which pollen dispersal is extensive and male-sterility systems are absent or unreliable. □

Alan J. Gray and Alan F. Raybould are at the Institute of Terrestrial Ecology, Furzebrook Research Station, Furzebrook Road, Wareham, Dorset BH20 5AS, UK.

e-mail: a.gray@ite.ac.uk

1. Daniell, H., Datta, R., Varma, S., Gray, S. & Lee, S.-B. *Nature Biotechnol.* **16**, 345–348 (1998).
2. Shah, D. M. *et al. Science* **233**, 478–481 (1986).
3. <http://www.olis.oecd.org/biotrack.nsf>
4. Hill, J. E. *Brighton Crop Prot. Conf.—Weeds* 1027–1034 (1995).
5. Kling, J. *Science* **274**, 180–181 (1996).
6. Raybould, A. F. & Gray, A. J. *Trends Ecol. Evol.* **9**, 85–89 (1994).
7. Bradshaw, L. D., Padgett, S. R., Kimball, S. L. & Wells, B. H. *Weed Technol.* **11**, 189–198 (1997).
8. Cousens, R. & Mortimer, M. *Dynamics of Weed Populations* (Cambridge Univ. Press, 1995).
9. Reboud, X. & Zeyl, C. *Heredity* **72**, 132–140 (1994).
10. Crawley, M. J. & Brown, S. L. *Proc. R. Soc. Lond. B* **259**, 49–54 (1995).
11. Charters, Y. M., Robertson, A., Wilkinson, M. J. & Ramsay, G. *Theor. Appl. Genet.* **92**, 442–447 (1996).
12. Boudry, P., Mörchen, M., Saumitou-Laprade, P., Vernet, P. & van Dijk, H. *Theor. Appl. Genet.* **87**, 471–478 (1993).

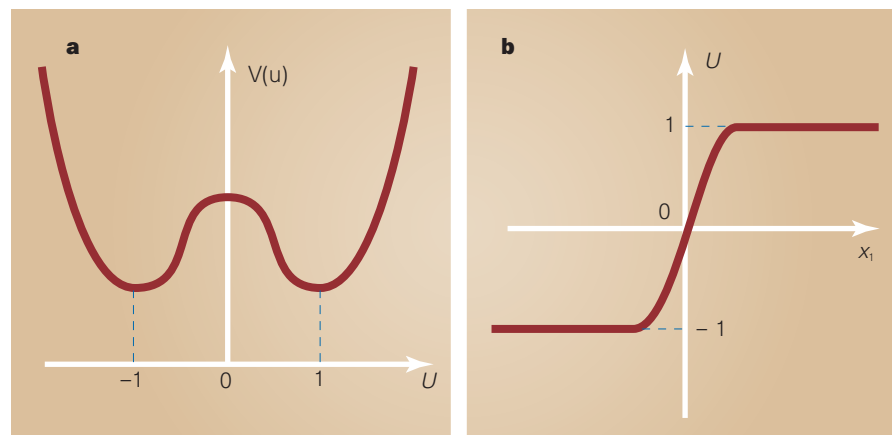


Figure 1 Potential and phase transitions. A simple system has two phases, characterized by some variable  $u$  which can have two stable values (defined by minima in the potential curve; a). Where two phases meet, there is a transition region (b), whose characteristics are completely determined by a few physical assumptions.



Figure 2 Don't try this in seven dimensions.

All his conjectures arise from physical considerations. One is concerned with phase transitions — melting ice, for instance. De Giorgi proposed a very simple model (Fig. 1a): a variable  $u$  describes the internal state of some matter, and a potential  $V(u)$  has two local minima, at  $-1$  and  $1$ , each corresponding to one phase. If  $u = -1$ , matter is in phase 1, if  $u = 1$ , matter is in phase 2, and for values between  $1$  and  $-1$ , matter is in some intermediate state, as we would expect, for instance, in a transition layer between the two phases.

But how does this intermediate state vary across the transition layer? On the one hand, it should vary smoothly: that is,  $u$  should be a smooth function of the position  $x = (x_1, x_2)$  (two dimensions) or  $x = (x_1, x_2, x_3)$  (three dimensions). On the other hand, physical intuition leads us to believe that the interface between the two phases should be sharp, that is, the transition layer should be narrow (Fig. 1b). To lend mathematical rigour to this intuition, one can start from the idea that the actual shape of the function  $u(x)$  should be obtained by minimizing some kind of overall energy, where the steepness (first derivative) of  $u$  comes into play.

Scaling the whole situation down to fit the width of the transition layer, de Giorgi was led to the following conjecture: if  $u(x)$  is a solution of  $\Delta^2 u - V'(u) = 0$ , defined over the whole plane or space, satisfying  $-1 < u < 1$ ,  $u/x_1 > 0$  (so  $u$  goes through zero at  $x_1 = 0$ , and nowhere else), and such that  $u$  goes to  $1$  or  $-1$  as  $x_1$  goes to infinity in the positive or negative direction; then  $u(x) = g(x_1)$ . In other words, the lines of constant  $u$  must be straight lines (in two dimensions) or planes (in three), and the state of matter at every point depends only on its distance from  $x_1 = 0$ .

The ansatz  $u(x) = g(x_1)$  clearly yields a solution. The point of the de Giorgi conjecture is that it is the *only* solution — the

physical model, although very naive, leads to a fully determined situation.

Besides the physical interpretation, the result is mathematically significant. The equation differs from the classical, linear Laplace equation ( $\Delta^2 u = 0$ , which describes, for example, electric fields outside regions of charge) only in its lower-order terms, yet its behaviour is completely different. Any bounded solution of the two-dimensional Laplace equation must be constant, and here, for very large  $x_1$ , the nonlinear equation is well approximated by the Laplace equation, with the corresponding constant being  $-1$  or  $1$ . But in the middle, the solution has to rise smoothly from one to the other, and the nonlinear terms somehow enable it to.

The de Giorgi conjecture has a long history, and the Ghoussoub–Gui proof is only the last step in its solution<sup>2–4</sup>. The technique they have developed is variational, that is, it relies on minimizing the energy integrals associated with certain equations. They first prove results concerning the eigenvalues of certain linear Schrödinger operators  $L = \Delta^2 u - f$  with bounded potential  $f$ , and they then use these results to prove the de Giorgi conjecture.

These intermediate results are interesting in themselves. Berestycki, Caffarelli and Nirenberg<sup>5</sup> proved last year that if  $L(u) = 0$  has a bounded solution  $u$  that changes sign, then  $L$  has negative-energy bound states, provided the underlying dimension is two or three. That raises the question, what happens in higher dimensions? Ghoussoub and Gui have now proved that this result is false if the underlying dimension is seven or higher. So seven-dimensional ice might have a very different melting behaviour (Fig. 2).

This dependence on the dimension, although not very intuitive, has often arisen in recent years in the field of PDEs. One of the most famous results is again connected with de Giorgi. It deals with functions whose



#### 100 YEARS AGO

"Balnibarbian Glumtrap Rhyme" –  
G. M. Minchin

Distant scintillating star,  
Shall I tell you what you are?  
Nay, for I can merely know  
What you were some years ago.

For, the rays that reach me here  
May have left your photosphere  
Ere the fight of Waterloo –  
Ere the pterodactyl flew!

Many stars have passed away  
Since your æther-shaking ray  
On its lengthy journey sped –  
So that you, perhaps, are dead!

Smashed in some tremendous war  
With another mighty star –  
You and all your planets just  
Scattered into cosmic dust!

Strange, if you have vanished quite,  
That we still behold your light,  
Playing for so long a time  
Some celestial pantomime!

But, supposing all is well,  
What you're made of, can I tell?  
Yes, 'twill be an easy task  
If my spectroscope I ask.  
From *Nature* 14 April 1898.

#### 50 YEARS AGO

In *Nature* of January 3, a review entitled "Atomic Energy" by Prof. J. A. Crowther referred to the progress made "since the spontaneous release of atomic energy fogged a packet of photographic material in the drawer of Becquerel's desk". This sentence seems to imply that the discovery of radioactivity by Henri Becquerel, my father, in 1896, was fortuitous – which is incorrect. He intentionally subjected a photographic plate – not a "packet" – to the action of salts of uranium. His first idea was that phosphorescence of uranyl salts might be accompanied by a radiation similar to X-rays, which had just been discovered. – Jean Becquerel  
From *Nature* 17 April 1948.

Many more abstracts like these can be found in *A Bedside Nature: Genius and Eccentricity in Science, 1869–1953*, a 266-page book edited by Walter Gratzer. Contact David Plant.  
e-mail: subscriptions@nature.com



graphs are minimal surfaces; in other words, functions  $f(x_1, \dots, x_n)$  for which the surface  $y = f(x_1, \dots, x_n)$  minimizes the area between all small closed curves drawn on it. Sergei Bernstein wondered, must such functions be linear — in other words, must their graphs be hyperplanes? The answer is yes if the underlying dimension is two (probably known to Bernstein himself, in the 1930s), three (de Giorgi), four (Almgren), five, six and seven (Simons). But Bombieri, de Giorgi and Giusti proved in 1968 that it is no longer the case in dimension eight or higher, prompting the question: “How is it possible that a theorem of analysis can be true for functions of seven variables and false for functions of eight variables?”

We are no longer surprised. Indeed, the intervening years have brought us a host of other important results showing that geom-

etry depends on the underlying dimension in ways that were not suspected before. The result of Bombieri, de Giorgi and Giusti tells us that there is something about eight-dimensional spaces that makes them inherently different from seven-dimensional ones. We now know similar things about dimensions four and three, which brings us deep into the structure of space-time — but that is another story. □

Ivar Ekeland is at the University of Paris-Dauphine, 75775 Paris, CEDEX 16, France.

e-mail: [ivar.ekeland@dauphine.fr](mailto:ivar.ekeland@dauphine.fr)

1. Ghoussoub, N. & Gui, C. *Mathematische ann.* (submitted).
2. Modica, L. & Mortola, S. *Boll. Un. Matematica Ital.* **5**, 614–622 (1980).
3. Modica, L. *Commun. Pure Appl. Math.* **38**, 679–684 (1985).
4. Caffarelli, L., Garofalo, N. & Segala, F. *Commun. Pure Appl. Math.* **47**, 1457–1473 (1994).
5. Berestycki, H., Caffarelli, L. & Nirenberg, L. *Annali della Scuola Normale Superiore di Pisa* (in the press).

## Developmental biology

# Pulling the fly's leg

Ginés Morata and Ernesto Sánchez-Herrero

Since Bridges isolated the first *bithorax* allele in *Drosophila* 75 years ago, such homeotic (Hox in modern parlance) genes have been at the centre of modern thinking on the genetic control of development. The Hox genes are conserved in evolution, suggesting that they perform similar functions in establishing the body plans of multicellular animals. Specifically, these genes ‘select’ the developmental pathways that are followed by cells, and, when mutated, produce impressive phenotypes in which one organ or segment is changed into another. For example, the spectacular transformation of antennae into legs that is seen in *Antennapedia* (*Antp*) mutant flies (Fig. 1) has been one of the models for the specificity of Hox genes in regulating developmental alternatives. Genetic and molecular studies indicate that the main function of *Antp* in adult flies is to specify leg development. But

on page 723 of this issue, Casares and Mann<sup>1</sup> report that *Antp* is not required to make a leg — instead, it prevents the functioning of *homothorax* (*hth*), a homeobox gene that is responsible for antennal development.

First a look back at the *Antp* story. The antenna-to-leg transformation of the dominant *Antp* mutants had been known for a long time, but, in 1981, Struhl<sup>2</sup> provided the critical results for a model of *Antp* function. By making revertants of a dominant *Antp*<sup>NS</sup> allele, he isolated recessive mutations that lacked *Antp* activity. In adult flies, these mutations produced a transformation of the second leg into an antenna — just the reverse of that seen in the dominant mutations — with no effect on the normal antennae. He concluded that *Antp* promotes development of the mesothoracic (second) leg, as opposed to antennal development, and suggested that *Antp* does this by suppressing a hypothetical

‘antennal-determining’ gene (or genes) that normally operates in the antenna but not in the leg.

In the absence of the two other Hox genes that are involved in leg patterning (*Sex comb reduced*, *Scr*, and *Ultrabithorax*, *Ubx*), *Antp* prevents antennal development in all three types of leg<sup>3</sup>. Struhl proposed that the phenotype of the dominant *Antp* mutations would result from inappropriate activity of the *Antp* product in the antenna, leading to repression of the antennal-determining genes. This was supported by data, mainly from Walter Gehring's laboratory, showing that the antenna-to-leg transformations are associated with the presence of the *Antp* protein in the antennae<sup>4,5</sup>.

All of this was well founded, but questions remained. First, the postulated antennal-determining genes were yet to be found. Second, unlike all other Hox genes, *Antp* controls different developmental alternatives in embryonic and adult cells: in *Antp*<sup>−</sup> embryos the mesothoracic segment develops like the first segment, prothorax<sup>6</sup>; in adults, however, *Antp*<sup>−</sup> mesothoracic leg cells differentiate into antennal structures. Third, during insect evolution the Hox genes are thought<sup>7</sup> to have created morphological diversity by modifying an ancestral thoracic-like segment. This view is supported by the phenotype of *Scr*<sup>−</sup>*Antp*<sup>−</sup>*Ubx*<sup>−</sup> embryos, in which all thoracic segments lack Hox activity and develop an identical thoracic ‘ground’ pattern<sup>8</sup>. However, *Scr*<sup>−</sup>*Antp*<sup>−</sup>*Ubx*<sup>−</sup> adult thoracic cells develop an antennal (cephalic) pattern<sup>3</sup>, and the improbable conclusion is that larvae have a thoracic ground pattern whereas adults have a cephalic one.

The paper by Casares and Mann<sup>1</sup> clarifies some of these issues. They identify the antennal-determining genes and also show how *Antp* suppresses their function. The genes turn out to be *extradenticle* (*exd*) and *hth* — two highly conserved and functionally related homeobox genes<sup>9,10</sup>. The Exd protein interacts with other Hox proteins, allowing them to bind Hox DNA targets *in vivo*<sup>11</sup>, and it is found in both the nucleus and the cytoplasm<sup>12,13</sup> (although only the nuclear form is functional). The *hth* gene positively regulates the function of *exd* by promoting the transport of Exd to the nucleus<sup>10</sup>. Because *hth* acts upstream of — and controls — *exd*, we refer to *hth* as being responsible for the developmental roles of both itself and *exd*.

Antennal cells that lack *exd* or *hth* differentiate leg patterns, suggesting that *hth* is involved in antennal development (ref. 1 and references therein). Casares and Mann now make this point more strongly by showing that ectopic expression of *Meis1*, a murine homologue that is functionally equivalent to *hth*<sup>10</sup>, transforms the analia (part of the cuticle surrounding the anus) into an antenna. This is a powerful result, because the ability of *hth* to trigger antennal development out of

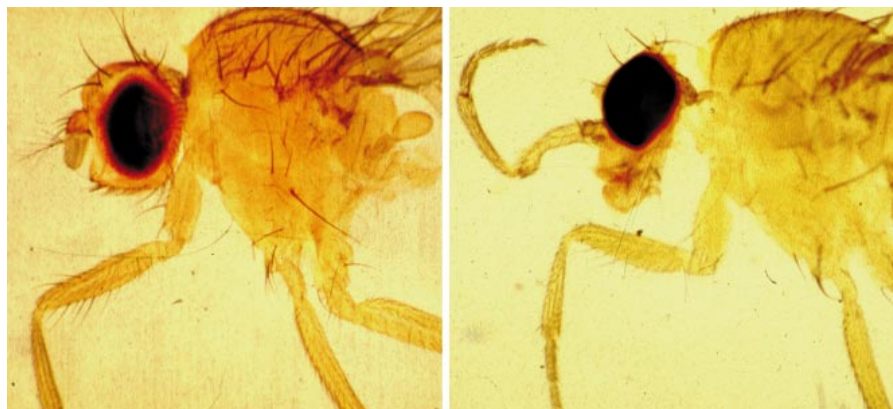
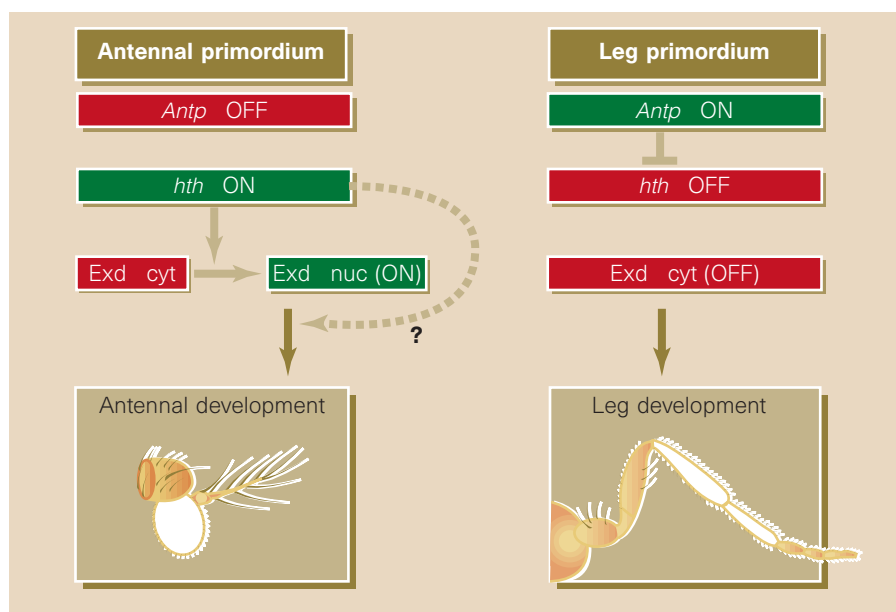


Figure 1 Head and thorax of a wild-type fly (left) and a dominant *Antennapedia* mutant (right). The only significant difference is the transformation, in the mutant fly, of the antenna into a leg. Although not visible here, the ectopic leg of the mutant fly has the bristle pattern and other features characteristic of the mesothoracic (second) leg.





**Figure 2** Interactions between *Antennapedia* (*Antp*), *homothorax* (*hth*) and *extradenticle* (*exd*) that determine the development of the antennal and leg primordia. Gene activity is marked in green and lack of activity in red. a, In the antennal primordium, *hth* translocates the Exd protein from the cytoplasm (cyt) to the nucleus (nuc), and antennal development follows. The dotted line with the question mark indicates the possibility that *hth* might have a direct role not mediated by *exd*, as proposed by Casares and Mann<sup>1</sup>. b, In the leg primordium the activity of *Antp* prevents transcription of *hth*, so Exd remains inactive in the cytoplasm and leg development ensues. Any mechanism that inactivates *hth* or prevents translocation of Exd (such as the presence of *Antp* protein in *Antp* dominant mutants) would transform the antenna into a leg.

the normal context indicates it acts as a master gene for the antenna. In the antennal disc (as in the anala) there are no active Hox genes to interact with, so the antennal expression of *hth* probably reflects a function that is independent of Exd/Hox interactions.

Now consider again the transformation to leg that is produced by *hth* or *exd* mutant cells in the antenna. This is the same phenotype as seen with dominant *Antp* mutants, but Casares and Mann show that, unlike in the *Antp* mutants, this leg develops without activity of *Antp*, *Scr* or *Ubx*. It follows that a leg can be generated without Hox activity, suggesting that the leg pathway is the ground state for ventral appendages. Thus the ground pattern for both larvae and adults is thoracic, removing one of the worries mentioned above. Nor does *Antp* 'select' for a specific leg pathway — it simply represses *hth* in the leg primordia, thereby blocking antennal development and allowing the development of legs by default (Fig. 2). This supports Struhl's model<sup>3</sup> that *Antp* promotes a ground (mesothoracic) pattern by repressing cephalic genes. This basal pattern is modified towards prothoracic (first leg) by *Scr* or metathoracic (third leg) by *Ubx* in their respective primordia.

The downregulation of *hth* by *Antp* explains the phenotype of the dominant *Antp* mutants as being due to *hth* repression. It also explains the ability of other Hox genes such as *Ubx*, *abd-A*, *Abd-B* and even the mouse *Hoxd-10* to induce the transformation of antennae into legs<sup>14,15</sup>. These genes prevent the nuclear translocation of Exd<sup>14</sup> (most likely through *hth* repression), so the antennal to leg transformations are probably nonspecific and caused by a property that is common to *Antp* and other Hox proteins.

During embryogenesis, in contrast with leg development, *Antp* selects for a specific developmental pathway. Loss-of-function mutations<sup>6,8</sup> and experiments to induce

ectopic expression<sup>16</sup> show that *Antp* determines the larval mesothoracic pattern — a function that is clearly distinct from the other Hox genes<sup>8</sup>. Why legs should be different is not clear, but different Hox genes have similar effects on appendages<sup>15</sup>, possibly because these appendages have no *hth* activity, without which the Hox genes lack specificity. □

Ginés Morata and Ernesto Sánchez-Herrero are at the Centro de Biología Molecular, CSIC-UAM Universidad Autónoma de Madrid, 28049 Madrid, Spain.

e-mail: gmorata@cbm.uam.es

1. Casares, F. & Mann, R. S. *Nature* **392**, 723–726 (1998).
2. Struhl, G. *Nature* **292**, 635–638 (1981).

3. Struhl, G. *Proc. Natl. Acad. Sci. USA* **79**, 7380–7384 (1982).
4. Schneuwly, S., Kuroiwa, A. & Gehring, W. *EMBO J.* **6**, 201–206 (1987).
5. Schneuwly, S., Klemenz, R. & Gehring, W. *Nature* **325**, 816–818 (1987).
6. Wakimoto, B. T. & Kaufman, T. C. *Dev. Biol.* **81**, 51–64 (1981).
7. Akam, M. *et al. Development (Suppl.)* 209–215 (1994).
8. Struhl, G. *J. Embryol. Exp. Morph.* **76**, 297–331 (1983).
9. Rauskolb, C., Peifer, M. & Wieschaus, E. *Cell* **74**, 1101–1112 (1993).
10. Rieckhof, G., Casares, F., Ryoo, H. D., Abu-Shaar, M. & Mann, R. S. *Cell* **91**, 171–183 (1997).
11. Mann, R. S. & Chan, S.-K. *Trends Genet.* **12**, 258–262 (1996).
12. Mann, R. S. & Abu-Shaar, M. *Nature* **383**, 630–633 (1996).
13. Aspland, S. E. & White, R. A. H. *Development* **124**, 741–747 (1997).
14. Azpiazu, N. & Morata, G. *Genes Dev.* **12**, 261–273 (1998).
15. Casares, F., Calleja, M. & Sánchez-Herrero, E. *EMBO J.* **15**, 3934–3942 (1996).
16. Gibson, G. & Gehring, W. *Development* **102**, 657–675 (1988).

## Condensed-matter physics

# The noise is the signal

Rolf Landauer

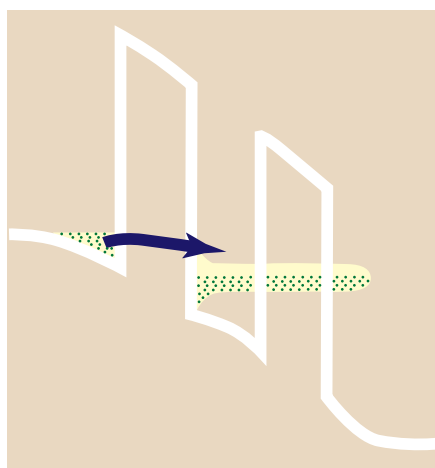
Noise is not only a hindrance to signal detection. Advances in measurement techniques mean that it can now be used to probe the kinetics of electrons. That is because interactions between electrons can regulate their relative motion, and so reduce noise. These interactions include not only the Coulomb repulsion, but also the Pauli exclusion principle, which prevents two electrons from behaving alike. Several new noise measurements<sup>1–4</sup> in different systems have produced insights into how these interactions affect electronic behaviour.

Eighty years ago, Schottky described the irregular 'patter' of electrons that cross a vacuum diode independently of one another, and called it shot noise. The mean-squared shot-noise current in a frequency range  $\Delta f$  is given by  $2eI\Delta f$ , where  $e$  is the charge on the electron, and  $I$  is the current. Some 20 years after Schottky, the case was described in which enough electrons are emitted from the cathode to repel each other and thereby

regulate and reduce the fluctuations.

The new investigations<sup>1–4</sup> were prompted by precise measurements of noise at quantum point contacts. These are narrow passages of controllable width whose conductance tends to be quantized<sup>5</sup> in multiples of  $e^2/\pi h$ . Experiments<sup>6</sup> at CEA Saclay found that the noise corresponds with what is predicted by the simplest existing models of electrons that move independently (except for the effects of the Pauli exclusion principle). This lack of evidence for interactions is not surprising, as the group's two-dimensional semiconductor structure does not have many electrons in the region that matters, so Coulomb interactions are small. But a Yale group<sup>1</sup> has made high-frequency noise measurements on a thin, short, narrow, gold strip with a high electron density, and again the theory that allows only for the Pauli principle accounts remarkably well for the results.

Theorists expect interactions to yield



**Figure 1 Noise with feedback.** Two potential barriers contain a bound state, whose energy overlaps that of electrons tunnelling into the well from the left. The shot noise is increased by positive feedback: excess electrons tunnelling in from the left raise the potential energy of the well, and so increase the overlap and allow still more tunnelling.

more complex beasts, but in practice they must be sought in more exotic country. In the fractional quantum Hall effect, for example, theory predicts the existence of pseudoparticles, collective electronic states with fractional electric charge. Indeed, two experiments<sup>7,8</sup> have now seen shot noise reduced from what would be expected if the carriers had the same charge as independently moving electrons, demonstrating the presence of carriers with an effective charge of  $e/3$ .

Noise increases have also been observed, caused by Coulomb interactions. A group from Pisa<sup>3</sup> measured shot noise in current flow through two successive tunnelling barriers (Fig. 1). A bound state between the two barriers has an energy width that overlaps with the conduction band. If a fluctuation

allows more electrons to enter the trap, the energy of trapped electrons rises, increasing the overlap and allowing more tunnelling. The shot noise is strongly increased by this positive feedback. This is the equivalent of the well-known critical fluctuations in thermal equilibrium, which appear as a state approaches instability. In this circuit the capacitance between the region containing the bound state and the two electrodes is shunted by the differential conductance of the two barriers in parallel. Instability of the voltage division between the two barriers arises when that net conductance becomes negative. Thus, the instability turns up when the magnitude of the negative differential conductance of the left barrier approaches that of the positive conductance of the right barrier. Such critical fluctuations in circuits were predicted decades ago<sup>9</sup>.

The Pauli principle alone can suppress noise, as the quantum-point-contact experiments<sup>6</sup> demonstrated. Another demonstration of this comes from a Stanford group<sup>2</sup>, who created the electronic equivalent of a half-silvered mirror (Fig. 2) by adjusting the height of a barrier that can transmit or reflect electrons. When the incident stream has all possible electronic states filled, it is regular and noiseless. The emerging streams are only half-filled with irregularly occupied wave-packet states, and so are noisy, as predicted<sup>10</sup>. When another fully occupied stream of incident electrons is added, now supplying enough electrons to fill all outgoing states, the noise in the emerging streams is greatly reduced.

As illustrated in Fig. 2, the Pauli principle allots at most one electron to each successive wave packet, and this can regulate the flow and reduce noise. Consider a conductor that is short enough to scatter traversing electrons in a coherent quantum-mechanical way, but long enough to transmit only a small proportion of the incident carriers,

with most scattered back to the incident interface. At the output end we might expect few of the possible emerging electronic states to be occupied, and therefore to see little of the exclusion-principle noise reduction. This turns out to be wrong. Interference makes electron transmission through the sample highly non-uniform, like a laser speckle pattern, and this allows the Pauli principle to be effective, yielding a noise that is  $1/3$  of the classical shot noise<sup>11</sup>.

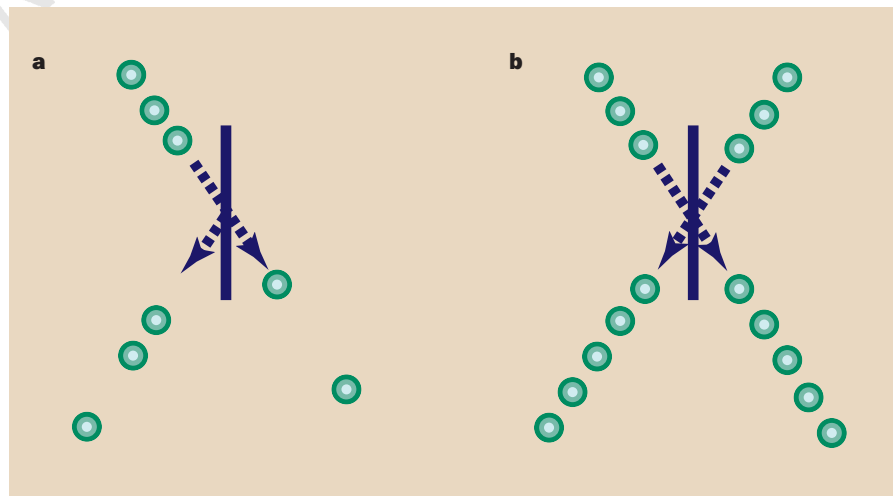
This factor of  $1/3$  applies in a wider range of circumstances, including the case where electrons are scattered incoherently, but elastically. I have claimed<sup>12</sup> that this occurrence of  $1/3$  in the two cases is just a numerical coincidence; that there is no common physical mechanism. Remarkably, now, the same factor has turned up in another, very different, context<sup>4</sup>. A computer simulation shows that it also emerges where there are too few electrons for the Pauli principle to matter, but where the current-carrying electrons make an additional charge injected into a previously neutral material, they are scattered elastically and there is strong Coulomb repulsion between them. But the supposed universality of  $1/3$  is spoiled somewhat by the fact that a two-dimensional velocity distribution in this case would instead yield a noise reduction by a factor of two. So totally different physical processes can lead to the same factor-of-three reduction.

Noise measurements are only just emerging as a useful quantitative probe of electron kinetics, and many questions remain. A particularly intriguing proposal<sup>13</sup> concerns two unconnected conducting paths: the paths overlap at two separate places in space, but electrons from one path cannot cross into the other path. A magnetic flux in the space enclosed between the two points of overlap will have no effect on the current flow in either path, yet it should affect the correlation in noise between the two paths. □

Rolf Landauer is at the IBM Thomas J. Watson Research Center, PO Box 218, Yorktown Heights, New York 10598, USA.

e-mail: landauer@us.ibm.com

- Schoelkopf, R. J., Burke, P. J., Kozhevnikov, A. A., Prober, D. E. & Rooks, M. S. *Phys. Rev. Lett.* **78**, 3370–3373 (1997).
- Liu, R. C., Odom, B., Yamamoto, Y. & Tarucha, S. *Nature* **391**, 263–265 (1998).
- Iannaccone, G., Lombardi G., Macucci, A. & Pwellegrini, B. *Phys. Rev. Lett.* **80**, 1054–1057 (1998).
- González, T. et al. *Phys. Rev. Lett.* **80**, 2901–2904 (1998).
- van Houten, H. & Beenakker, C. *Physics Today* 22–27 (July 1996).
- Kumar, A., Saminadayar, L., Glattli, D. C., Jin, Y. & Etienne, B. *Phys. Rev. Lett.* **76**, 2778–2781 (1996).
- de-Picciotto, R. et al. *Nature* **389**, 162–164 (1997).
- Saminadayar, L., Glattli, D. C., Jin, Y. & Etienne, B. *Phys. Rev. Lett.* **79**, 2526–2529 (1997).
- Pytte, E. & Thomas, H. *Phys. Rev.* **179**, 431–443 (1969).
- Martin, Th. & Landauer, R. *Phys. Rev. B* **45**, 1742–1755 (1992).
- Beenakker, C. J. W. & Büttiker, M. *Phys. Rev. B* **46**, 1889–1892 (1992).
- Landauer, R. *Physica B* **227**, 156–160 (1996).
- Büttiker, M. *Phys. Rev. Lett.* **68**, 843–846 (1992).



**Figure 2 Tightly packed electrons.** a, A stream of electrons is split into equal parts at the barrier. All the possible incident states are filled with an electron, yielding a regular stream. The emerging half-filled streams are more random, and therefore noisy. b, When another stream of incident electrons is added, both emerging streams are full, and less noisy.

Animal navigation

# Ants match as they march

Mandyam V. Srinivasan

Behavioural biologists have long been intrigued by the strategies that animals use to find their way to a feeding site and return home. Hymenopterans, given their miniature brains, are particularly impressive navigators. Bees, wasps and certain species of ant rely heavily on vision for navigation and are very adept at using prominent landmarks, when available, as aids. The paper by Judd and Collett on page 710 of this issue<sup>1</sup> provides new insights into how landmarks are memorized and used.

Nearly 20 years ago, Wehner and R  ber<sup>2</sup> and Cartwright and Collett<sup>3</sup> suggested how an insect, when close to its goal, might make use of the surrounding landmarks to pinpoint that goal. Upon its first visit to the goal, the insect takes a 'snapshot' of the panorama as seen from the goal. On its subsequent visits the insect finds the goal by positioning itself so that the current image of the environment matches the stored snapshot. If the image of a particular landmark is smaller than in the memorized snapshot, the insect moves closer to that landmark to enlarge its image; if the image is larger, the insect moves away. In other words, the insect homes in on the goal by matching the current view to a 'retinotopically learned' image.

This strategy works well when the insect is in the vicinity of the goal, but runs into difficulties when the insect is far away. There,

the sizes and arrangement of the landmarks' images on the retina may be very different from what the insect experiences near the goal. Worse yet, at a remote location the insect is likely to be surrounded by a different set of landmarks, and those near the goal may not even be visible (Fig. 1). It has been speculated that navigation over long routes might be accomplished by acquiring a sequence of

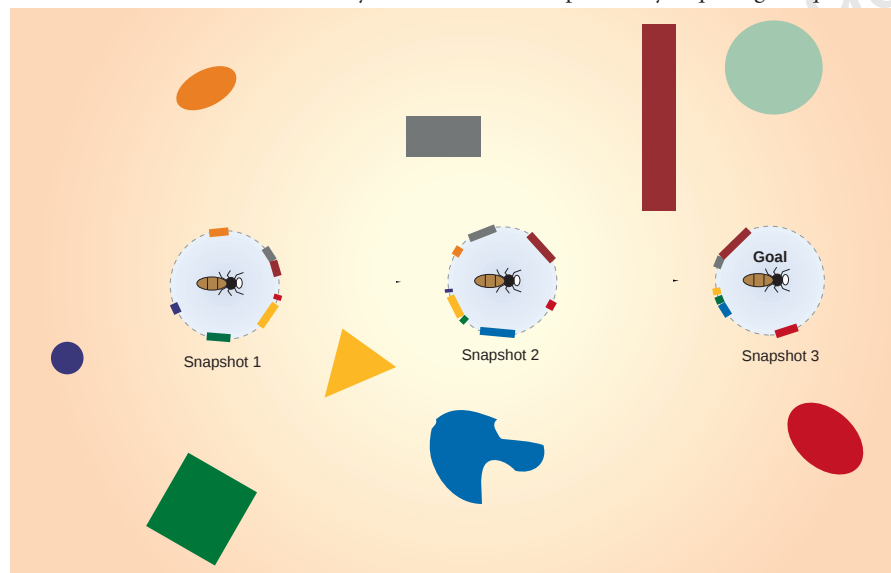


Figure 1 Illustration of how an ant might navigate to its goal (say, a food source) by using a sequence of 'snapshots' of the environment acquired during the learning phase after it first visited that goal. From their experiments, Judd and Collett<sup>1</sup> propose that the snapshots are acquired during so-called inspection runs — when, after leaving the goal for the first time, the ants frequently turn around and make short runs back towards it.

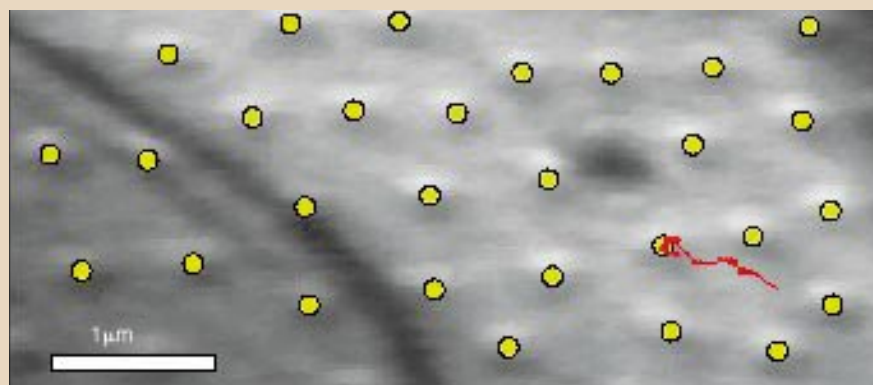
## Superconductors

### Cold cousin of the colloids

Some superconductors admit magnetic fields in thin flux lines, or vortices. Any electrical current exerts a force on the vortices, and if that force is enough to make them move, they dissipate energy. In other words, the vortices cause electrical resistance, even though the material around them remains superconducting.

The high-temperature superconductors are all of this type, and the misbehaviour of magnetic flux quanta severely limits their usefulness by reducing the maximum current and magnetic field that they can bear. The obvious way to solve this problem is to pin the vortices down: if they can't move, they can't waste energy.

Vortices are naturally pinned in place by defects in the materials, but weakly. To investigate this 'pinscape', in the hope of understanding how to make the pinning more effective, a Japanese group led by Akira Tonomura made high-resolution videos of vortex motion (a 'still' is shown here). Looking at a piece of niobium at 4.5 K with an electron microscope, they watched flux lines stagger from one



pinning site to another, propelled by a gradient in the magnetic field. The red thread marks such a path.

Watching these moving pictures, David Grier of the University of Chicago was reminded of Brownian motion in colloids, and realized that the depth of the pinning-site potential wells has the same randomizing effect as the thermal energy of colloid particles — or of any liquid particle. He and his colleagues have now adapted their data-analysis techniques to

the vortex movies, and use them to measure the strength of vortex–vortex interactions and the distribution of pinning potentials (C.-H. Sow *et al.* *Phys. Rev. Lett.* 80, 2693–2696; 1998). These numbers are already known for the simple low-temperature superconductor niobium, but the method could provide unique information about the more complicated high-temperature superconductors.

Stephen Battersby



snapshots on the journey, and steering the correct route by matching each snapshot in succession<sup>4</sup>.

Judd and Collett<sup>1</sup> now provide evidence that wood ants may indeed adopt such a strategy. They trained ants to find food behind a specific landmark, and video-filmed the ants' paths as they walked to the correct landmark and returned to their nest. Analysis of the paths revealed that the ants were operating not just on a single snapshot of the landmark, taken very close to the landmark where the food was located, but on a series of snapshots acquired at various points along the route.

How and when are the snapshots acquired? During the learning phase, ants leaving for home after collecting their food frequently turn back and make short runs towards the goal. Judd and Collett speculate that the snapshots may be acquired during these 'inspection' runs. The behaviour of learning ants leaving the feeding site is reminiscent, in some ways, of how bees 'turn back and look' when they leave a newly discovered feeder<sup>5</sup>, or how wasps departing from their nests inspect the environment in the nest's vicinity<sup>6</sup>. The precise function of these striking aerial manoeuvres remains a mystery, although there is some evidence that they may be used to extract information about the three-dimensional structure of the immediate environment<sup>5,6</sup>.

In Judd and Collett's experiments, the ants did not have to deal with a visual environment that changed drastically as they proceeded from the nest to the food site. The food-bearing landmark acted as a beacon which guided them towards it, and produced a steadily growing image on the retina. For this task, therefore, the ants might well have managed with a single snapshot acquired at the goal. Nevertheless, the data indicate that they acquired not one, but a series of snapshots at different distances, and recalled them in the appropriate sequence. This finding reflects the potential capacity of the ant's navigational machinery to deal with visual environments such as that shown in Fig. 1, which are complex and rapidly changing, and which provide no beacon at which to aim.

One way in which a sequence of snapshots could be used in such circumstances would be to associate each snapshot with a vector which specifies how far and in what direction the ant should move to arrive at the location at which the next snapshot was acquired, as shown by the arrows in Fig. 1. Another strategy would be to dispense with the motion vectors and, instead, to acquire snapshots at closer intervals than those illustrated in the figure. If the rate of snapshot acquisition is sufficiently high, the insect could travel from each snapshot location to the next simply by recalling the next snapshot from memory and moving so as to

match the viewed scene to it. This second possibility is supported by Judd and Collett's observation that the wood ants acquire snapshots more frequently when the image of the environment changes rapidly.

Do insects use 'photographic' representations for tasks other than navigation? One would surmise that a miniature brain with restricted memory capacity would favour compact, non-photographic representations whenever feasible. Bees seem to learn certain properties of flowers, for instance the spatial arrangement of the colours, in a photographic way<sup>7</sup>. But they also seem to learn other attributes of the same flowers — such as the dominant orientation of the constituent contours, and the radial or bilateral symmetry of the shape — in an abstract fashion<sup>8–10</sup>, and recognize these general attributes in other objects that they have never previously encountered.

Future research will sort out the extent to which insects can be described as photographers taking snapshots, or as abstract artists capturing the essence of what they see. In the

meantime, it is interesting to note that engineers have independently discovered the 'snapshot principle' and are applying it to the design of algorithms for autonomously navigating robots<sup>11</sup>. □

Mandyam V. Srinivasan is at the Centre for Visual Sciences, Research School of Biological Sciences, Australian National University, PO Box 475, Canberra, ACT 2601, Australia.

e-mail: M.Srinivasan@anu.edu.au

1. Judd, S. P. D. & Collett, T. S. *Nature* **392**, 710–714 (1998).
2. Wehner, R. & R ber, F. *Experientia* **35**, 1569–1571 (1979).
3. Cartwright, B. A. & Collett, T. S. *J. Comp. Physiol.* **151**, 521–543 (1983).
4. Wehner, R., Michel, B. & Antonson, P. *J. Exp. Biol.* **199**, 129–140 (1996).
5. Lehrer, M. *J. Exp. Biol.* **199**, 253–261 (1996).
6. Zeil, J., Kelber, A. & Voss, R. *J. Exp. Biol.* **199**, 245–252 (1996).
7. Wehner, R. in *Handbook of Sensory Physiology* Vol. VII/6C (ed. Autrum, H.) 287–616 (Springer, Berlin, 1981).
8. Giger, A. & Srinivasan, M. V. *J. Comp. Physiol. A* **176**, 791–795 (1995).
9. Horridge, G. A. *J. Insect Physiol.* **42**, 755–764 (1996).
10. Giurfa, M., Eichmann, B. & Menzel, R. *Nature* **382**, 458–461 (1996).
11. Hong, J., Tan, X., Pinette, B., Weiss, R. & Riseman, E. M. *Proc. IEEE Int. Conf. Robotics and Automation*, Sacramento, CA 620–625 (IEEE, Piscataway, NJ, 1991).

## Plant extinction

# Frondless ferns lie low to survive

Peter D. Moore

Why do we write to the newspapers to record the first cuckoo of spring, yet never the last cuckoo of the summer? The answer, of course, is that we cannot be sure that it was the last cuckoo until time has proved us correct. In the same way, it is difficult to be sure when we have witnessed an extinction — particularly in the case of a plant. When a plant disappears from a locality, there is always the possibility that it will emerge from some obscure and unobserved dormancy, perhaps decades later, to the predictable delight of conservationists. This has recently happened with a fern that was thought to be extinct over eastern Britain and through most of continental Europe. The story of the resurrection of the Killarney fern (*Trichomanes speciosum*) is reported by Rumsey, Jermy and Sheffield<sup>1</sup> in *Watsonia*, the journal of the Botanical Society of the British Isles.

The Killarney fern is particularly beautiful, with delicate, bottle-green, translucent fronds (Fig. 1, left; overleaf). It grows in deep shade, often in dark, wet, rocky ravines in oceanic locations. Obsessive Victorian fern-hunters collected this fern, and some of their specimens still survive in captivity a century later<sup>2</sup>. But, by the mid-nineteenth century, it had already been lost over much of its former range. Writing in 1867, the Victorian pteridologist E. J. Lowe recorded<sup>3</sup> that the Killarney fern had grown in abundance near Bingley in Yorkshire in 1758, but had

been reduced to one root by 1782. The sporophyte generation, which bears the conspicuous fronds that carry the spore-bearing structures, is now known from only a few localities in western parts of Britain and Ireland (including Killarney), Brittany, western Iberia and some of the Atlantic islands, including Cape Verde, the Canaries, the Azores and Madeira.

Ferns have a complex life cycle, involving a free-living, independent stage with a haploid constitution — the gametophyte. In the case of the Killarney fern, the gametophyte is perennial, and structurally very distinctive. It is filamentous, rhizoidal and highly branched, forming a green mat over the surface of wet rocks (Fig. 1, right). The gametophyte bears not only the sexual organs, but also asexual gemmae that can break off and form new gametophyte plants. These features help to identify the gametophyte of *Trichomanes*.

The gametophyte generation of most ferns is a short-lived, flat plate of cells, the sole function of which is to bear the sexual organs. But in certain genera, such as *Trichomanes*, it can form an independent and persistent plant that is more tolerant of desiccation and frost than are the sporophytes<sup>4</sup>. Moreover, in some North American species, gametophytes have been found in locations that are geographically isolated from populations of the sporophyte.

Over the past few years, there has been an intensive search to see whether the Killarney



Figure 1 Alive and well — the Killarney fern, *Trichomanes speciosum*. Left, the sporophyte generation. Right, the gametophyte generation, now discovered by Rumsey *et al.*<sup>1</sup> in areas where the fern was thought to be extinct.

fern, presumed extinct over its former eastern range, may survive in a gametophyte form in some of its former localities. Rumsey, Jermy and Sheffield now report<sup>1</sup> that the gametophyte has been found alive and well in eastern England (from Yorkshire to Kent), France, Germany, northern Italy and even east into the Czech Republic. The plant has evidently continued its existence in a more occult form, but a desiccation-resistant form which has proved more appropriate for the warmer, drier climate that developed after the Little Ice Age (ending around 1700).

Several lessons, and questions, arise from this discovery. The definition of extinction for a plant must cover all stages of its life cycle. Seeds from plants considered extinct in a locality have been known<sup>5</sup> to germinate from soil seed-banks, often due to physical disturbance, and have re-established an above-ground population. For example, the fen violet (*Viola persicifolia*) reappeared, following some clearance of scrub, after an absence of almost 70 years at a wetland site in eastern Britain.

The seed bank must, therefore, be regarded as part of the population of a higher plant species when considering its conservation and biogeographical status. But this is difficult to survey, and the status of plants that occupy ephemeral habitats — and, as a result, have seeds with long dormancies — is a particular headache. For instance, a rare subspecies of the perennial knawel (*Scleranthus perennis* ssp. *prostratus*), which is endemic to East Anglian heaths<sup>6</sup>, shows short-term disappearance but long-term persistence.

Some flowering plants (such as many orchids) may even spend part of their lives in a subterranean form, before emerging and

becoming autotrophic. Orchids are notoriously difficult to monitor in population terms, because juvenile plants indulge in symbiotic, saprophytic nutrition. Some orchids, such as the Western Australian *Rhizanthella gardneri*, rarely show themselves above ground<sup>7</sup>, and then only for flowering. Such plants could easily become extinct, unnoticed.

Another important issue is that of protection. In the case of the Killarney fern, Rumsey *et al.*<sup>1</sup> point out that strict legislation exists to protect the fern sporophyte, and that this should be regarded as covering the gametophyte. But it may be difficult to enforce such a law, because the gametophyte superficially resembles an algal felting — a green fuzz on a wet rock. Their survey indicates, however, both a widespread distribution and a degree of local abundance for the Killarney fern, so it may be less threatened than was feared. Moreover, its unappealing gametophyte may not need to be legally protected from collectors. The Killarney fern may one day come into its own in the east once more, and develop some fruiting fronds — but it may have to wait until the climate goes into reverse and we have another Little Ice Age. □

Peter D. Moore is in the Division of Life Sciences, King's College, Campden Hill Road, London W8 7AH, UK.

1. Rumsey, F. J., Jermy, A. C. & Sheffield, E. *Watsonia* 22, 1–19 (1998).
2. Page, C. N. *The Ferns of Britain and Ireland* 2nd edn (Cambridge Univ. Press, 1997).
3. Lowe, E. J. *Our Native Ferns* Vol. 2 (Groombridge, London, 1867).
4. Farrar, D. R. *Proc. R. Soc. Edinb.* 86B, 361–369 (1985).
5. Moore, P. D. *Nature* 303, 572 (1983).
6. Crompton, G. in *The Biological Aspects of Rare Plant Conservation* (ed. Synges, H.) 117–124 (Wiley, Chichester, 1981).
7. Given, D. R. *Principles and Practice of Plant Conservation* (Chapman & Hall, London, 1994).

## Daedalus

### Cleansing fire

The ramjet is an appealingly simple engine, with no moving parts. Its sheer forward velocity compresses the air entering its intake; fuel is burnt in it, and the resulting hot gas expands in the exhaust nozzle to generate thrust. Sadly, it only works at very high speeds. Daedalus once invented a more tractable variant. This was a ground-effect ramjet, or gramjet, which skimmed over a prepared smooth track. As it sped along, its sloping underside compressed the air beneath it against the track. Downward-firing burners raised its temperature, and the hot exhaust gases expanded against its upwardly sloping rear surface to drive it along. Ground-effect compression works at quite moderate speeds, and the upward pressure of entrainment and combustion levitated the craft in hovercraft fashion. Inevitably, however, the fierce unshielded flame beneath it tended to damage its track.

Daedalus now wants his gramjet to travel over water. He sees it as the ideal way of cleaning up oil slicks. Not only would the intense flame from the vehicle burn the oil; the heat of its combustion would help to drive it along. Indeed, a fresh, thick slick loaded with volatile hydrocarbons might even burn fiercely enough to drive the gramjet for free.

Sadly, most aged slicks are too thin and tenuous to burn as strongly as this. In any case, a thin slick is water-cooled too well to burn completely. But Daedalus is undismayed. He recalls that a flame passed over a polyethylene sheet oxidizes its surface (this makes it easier to glue). Even if a slick is only partly oxidized by flame, the resulting oxygenates and fatty acids will act as detergents, emulsifying and dispersing the unburnt portions. So DREADCO engineers are now devising the most chemically active flame for the slick-burning gramjet. Ozone will enhance its reactivity; an electric discharge through the flame will generate active free radicals; oxidants such as sodium peroxide will be sprayed in to encourage the combustion reaction and form detergent from its residues.

Thus a powerful new broom will sweep our polluted seas and waterways. Waterborne gramjets will zoom over oceanic slicks, around dirty harbours and along fouled industrial canals, leaving a sparkling wake behind them. Unless, of course, it is more profitable to keep the water thickly covered in oil, and run slick-fuelled gramjet services over it.

David Jones



# Parker's pieces

Some readers were bemused, and others amused, to find a series of images by Cornelia Parker in the pages of *Nature* last autumn. Her work invites us to let our imaginations run free.

## Martin Kemp

“What’s this bollocks doing in *Nature*?” The question, more rhetorical than inquisitive or grammatical, was loudly posed in the biology department’s tea room at Leicester University by a postgraduate faced with the enigmatic insertion in the journal last October of Cornelia Parker’s photographs of “Navel fluff” from a “sailor (while at sea)”, a “farmer” and an “architect” (*Nature* 389, 668; 1997). In truth, none of Parker’s magnified images of delicate residues insinuated by the editor in that or the three preceding issues looked visually out of place. The problem was one of the context for meaning.

Parker’s pieces do not rely on fixed meaning. They do not communicate with the studiously unambiguous parades of hypothesis, evidence, analysis and demonstration that is the aspiration of articles and letters in this journal.

**Rather they exploit a resonant series of visual and verbal associations to provide invitingly open fields for our imaginative exploration. Residues of memory – of the acts that have deposited what we now see, or of their pedigree of ownership – are central to her associative processes.**

The scored furrows in microphotographs of a record of the “Nutcracker Suite” once owned by Hitler and the lacquer residue etched from a master recording at the Beatles’ Abbey Road Studios (*Nature* 389, 335 & 548; 1997) exploit the familiar frisson of fame by association and also strike deeper — into such issues as sweet music for an evil listener and “negatives of sound” (as she calls them).

*Cold Dark Matter* exemplifies the voraciously encyclopedic reach of her associative creations, a reach that extends from childish play to the cosmos of science. To begin with, an inoffensive garden shed, illuminated from within and stuffed with acquired clut-



Parker’s *Cold Dark Matter: An Exploded View* (Shed and contents being blown up by the army), 1991.

ter, stood in an otherwise empty gallery. Then, courtesy of the British Army, the shed was blown apart in a centrifugal cascade of fractured planks and material possessions — a blasted plastic dinosaur, a boot rent asunder, a savaged hot-water bottle, a euphonium wrenched out of tune.... Systematically harvested, the debris has been reassembled in the gallery as a galaxy of fragments hanging on wires. Lit by an inner bright star, its shadows collide with the confining walls, roof and floor.

Parker is not illustrating a scientific concept. Rather she is realizing the metaphorical poetry inherent in science’s cool analyses and calculated acts of naming. The moment of catastrophic expansion is frozen in a state of uncertainty, its ‘implicate order’ (in the physicist David Bohm’s sense) suspended between total disintegration and potential

reconstitution. The ‘exploded view’ plays on the convention of technical illustration, virtually invented by Leonardo da Vinci, in which a mechanism is visually dismembered to allow the discerning and naming of the parts.

*Cold Dark Matter* orbits around themes of gathering and dispersal, expansion and contraction, explosion and implosion, inhaling and exhaling, falling and suspension, transformation and stasis, peace and violence, making and breaking, playful and sinister — alternative states existing simultaneously within the phenomenon, and available for oscillating perception by viewers who are willing to free themselves to participate in her mind games. □

Martin Kemp is in the Department of the History of Art, University of Oxford, 35 Beaumont Street, Oxford OX1 2PG, UK.





# Tree stem diameters fluctuate with tide

The diameter of tree stems growing under open and controlled conditions undergoes rhythmic fluctuations independent of daily periodic factors such as light, temperature and humidity. We find a strong correlation between these fluctuations and the timing and strength of tides. This correlation suggests that the Moon is influencing the flow of water between different parts of the trees.

Rhythmic variations of stem diameter occur both in young trees in an environment where temperature, humidity and photoperiod are controlled, and in stem sections that are disconnected from the root system and crown<sup>1</sup>. These cycles occur simultaneously in independent trees; continue for several months in isolated stem segments, as long as the cambium remains alive; and persist under continuous light and darkness, during the period of vegetative dormancy or under reversed photoperiod<sup>1</sup>. The cycle usually appears as a double-peaked wave with a period of about 25 hours.

The period and nature of the variation are reminiscent of the variations in level of the water table in boreholes, or of the flow of springs. Such variations are of the same nature as coastal tides, but with a much

reduced amplitude and inverted values. In fact, every point of the Earth is subject to calculable 'Earth-tides', which result principally from the interaction of the lunar rhythm (daily period, 24 hours 49 minutes; monthly period, 29.5 days) with the respective positions of Earth and Sun.

In trees, the states of germination and initial growth relate to the phases of the Moon<sup>2,3</sup>, and an overview of this phenomenon based on about 600 animal and plant species or groups has been published<sup>4</sup>. However, the effect of the daily lunar synodic rhythm (to which the tides correspond) on trees, or on vascular plants in general, remains unexplored. Research based on the hypothetical existence of such a phenomenon has been started<sup>5</sup>, and variations in the diameter of elder tree (*Alnus* sp.) stems have shown similarities to daily gravimetric curves (D. M., unpublished results).

We compared several extensometric curves, obtained between 1977 and 1991, with the inverse values of the calculated gravimetric curves for the location of the trees (tide values were supplied by the Hydrographic and Oceanographic Service of the French Navy, Brest/Institute of Geodesy and Photogrammetry, Swiss Federal Institute of Technology, Zurich). We saw a

systematic correlation, although in one case there was a phase shift.

For example, two young spruce trees (*Picea abies* L. Karst.) of about 150 cm in height, grown in independent containers between 17 and 20 July 1988 in an air-conditioned environment with controlled temperature and continuous darkness, showed simultaneous rhythmic variations in diameter, distinguishable from each other only by their amplitude (Fig. 1a). In particular, the double peaks were noticeable in each of the two series of measurements. Superimposing the curve of the Earth-tides (Fig. 1b) revealed them to be almost completely synchronous.

At this time of the year, a spruce in the open air undergoes a fall in transpiration flow daily at about 18:00–19:00 hours<sup>6</sup>. This emphasises the fact that the extensometric and gravimetric absolute minima coincide. We noticed in this case that the depression takes place at 10:00, 13:00 and about 14:00 hours in a controlled environment. This part of the experiment involves young trees that present no photosynthesis, as they are maintained in continuous darkness, and that have stomata that are closed, excluding the upward flow of sap due to transpiration<sup>7,8</sup>.

The relationship between trees and the daily lunar synodic rhythm is particularly obvious when growth conditions in a controlled environment reduce the interference of diurnal photoperiods<sup>9</sup> to a minimum. This phenomenon is also apparent on isolated stem sections, but only as long as the cambium is alive.

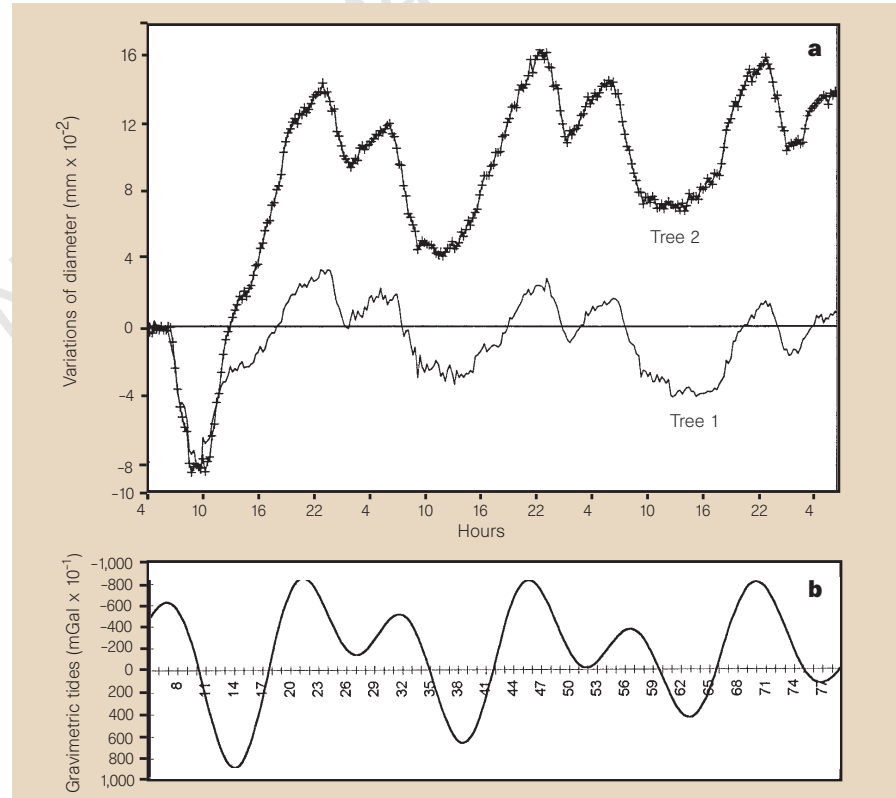
This, together with the fact that the water mass included in the measured volume is constant, indicates that a rhythmic, active and reversible flow could be involved from the symplast (the network of living cellular cytoplasm communicating through plasmodesmata) to the apoplast (the frame of dead cell walls, external to the plasmalemmas). The physiology of plants reveals that, for short distances, both systems contribute to the water transport process<sup>10,11</sup>. However, every short-term increase in diameter would correspond to a phase in which water accumulated in the cell walls. In wood technology this is called 'bound water', and leads to measurable 'swelling'.

**Ernst Zürcher**

Chair of Wood Sciences, Swiss Federal Institute of Technology, 8092 Zurich, Switzerland  
e-mail: zurcher@waho.ethz.ch

**Maria-Giulia Cantiani**

Department of Civil Engineering and Environmental Sciences, University of Trento, Mesiano di Povo, 38050 Trento, Italy



**Figure 1** Reversible diameter variations in spruce and gravimetric tides. **a**, Two independent young spruce trees under conditions of constant darkness and controlled temperature show simultaneous reversible diameter variations, which are strongly correlated (synchronous) with **b**, the gravimetric curve calculated for the same place (Florence, Italy) and period (17–20 July 1988).

Francesco Sorbetti-Guerri

Department of Agronomy and Forestry,  
University of Florence, Via S. Bonaventura 13,  
50145 Florence, Italy

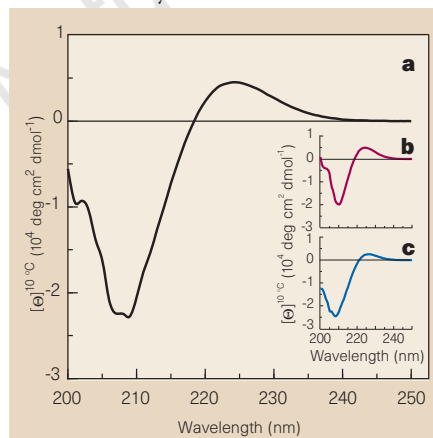
Denis Michel

Bureau d'Etudes et de Recherches Géologiques  
Appliquées, 1500 Avenue de la Pompiègne,  
34000 Montpellier, France

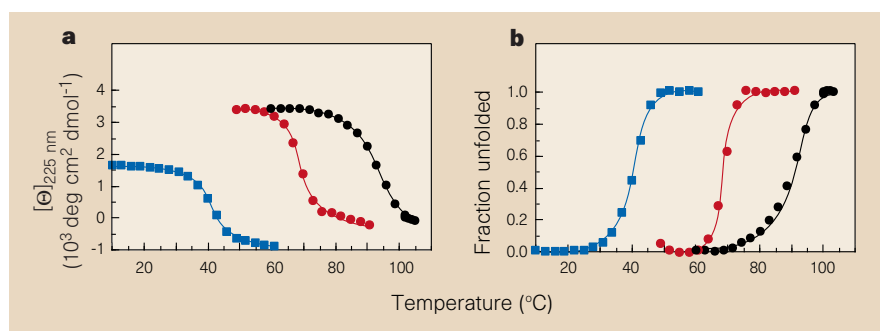
1. Cantiani, M., Cantiani, M.-G. & Sorbetti-Guerri, F. *Revue Forestière Française* **XLVI**, 349–358 (1994).
2. Zürcher, E. *Schweizerische Zeitschrift für Forstwesen / Journal Forestier Suisse* **143**, 951–966 (1992).
3. Zürcher, E. in *L'Arbre, Biologie et Développement. Troisième Colloque International. Montpellier, 11–15 Sept. 1995* (ed. Edelin, C.) 150–164 (Naturalia Monspenliensis, Montpellier, 1995).
4. Endres, K.-P. & Schäd, W. *Biologie des Mondes. Mondperiodik und Lebensrhythmen* (S. Hirzel Verlag, Stuttgart/Leipzig, 1997).
5. Michel, D. *Programme de Mission 'Radeau des Cimes', Automne 1991* (Projet BERGA-SUD, Univ. Montpellier internal document, 1991).
6. Larcher, W. *Ökologie der Pflanzen auf physiologischer Grundlage* 3rd edn (Ulmer, Stuttgart, 1980).
7. Lyr, H., Fiedler, H. J. & Tranquillini, W. (eds) *Physiologie und Ökologie der Gehölze* (Fischer, Jena/Stuttgart, 1992).
8. Whitehead, D. & Teskey, R. O. *Tree Physiol.* **15**, 245–251 (1995).
9. Thomas, B. & Vince-Prue, D. *Photoperiodism in Plants* 2nd edn (Academic, New York, 1997).
10. Strasburger, E. *Lehrbuch der Botanik* 32nd edn (eds von Denffer, D. et al.) (Fischer, Stuttgart/New York, 1983).
11. Raven, P. H., Evert, R. F. & Eichhorn, S. E. *Biology of Plants* 5th edn (Worth, New York, 1992).

## Code for collagen's stability deciphered

The most abundant protein in animals is collagen. In connective tissue, this protein is present as chains wound in tight triple helices which are organized into fibrils of great tensile strength and thermal stability<sup>1,2</sup>. We propose a new explanation for this stability.



**Figure 1** Circular dichroism spectra of collagen-related peptides. **a**, (ProFlpGly)<sub>10</sub>. **b**, (ProHypGly)<sub>10</sub>. **c**, (ProProGly)<sub>10</sub>. (ProFlpGly)<sub>10</sub> was synthesized by segment condensation of FmocProFlpGlyOH units on 2-chlorotrityl chloride resin. (ProHypGly)<sub>10</sub> and (ProProGly)<sub>10</sub> were supplied by Peptides International. Each 30-mer was >90% pure. Circular dichroism spectra were recorded in 50 mM acetic acid at 10 °C after incubation for 24 hours at 4 °C.



**Figure 2** Thermal denaturation of collagen-related triple helices. **a**, Raw data. **b**, Transformed data. Peptides (0.23 mM) were incubated in 50 mM acetic acid for 24 hours at 4 °C. Ellipticity at 225 nm was monitored by circular dichroism spectroscopy as the temperature was increased by increments of 3 °C with 5-minute equilibration. The values of  $T_m$  (°C), which is the temperature at the midpoint of the thermal transition curve, are: (ProProGly)<sub>10</sub> (blue),  $41 \pm 1$ ; (ProHypGly)<sub>10</sub> (red),  $69 \pm 1$ ; (ProFlpGly)<sub>10</sub> (black),  $91 \pm 1$ .

Each polypeptide chain of collagen is composed of repeats of the amino-acid sequence X–Y–Gly, where Gly represents glycine and X and Y are often proline (Pro) or 4(R)-hydroxyproline (Hyp) residues, respectively. The thermal stability of triple-helical collagen is enhanced by the hydroxyl group on the pyrrolidine ring of Hyp residues<sup>3</sup>. It is thought that this extra stability arises from hydrogen bonds mediated by a network of bridging water molecules<sup>4</sup>. Such water bridges have been detected by X-ray diffraction analysis of crystalline collagen<sup>5</sup>. They consist typically of two water molecules that link a Hyp side-chain of one strand to a main-chain carbonyl of another strand.

We suspected, for several reasons, that water bridges are unlikely to contribute significantly to collagen stability. First, the entropic cost of building and maintaining them would be enormous — the bridges would immobilize more than 500 water molecules per triple helix. Also, triple helices of (ProProGly)<sub>10</sub> and (ProHypGly)<sub>10</sub> are stable in methanol or propane-1,2-diol (ref. 6), and the Hyp residues enhance stability even in these anhydrous environments.

Electron-withdrawing groups can alter the attributes of molecules<sup>7</sup>. The inductive effect of the hydroxyl group of Hyp residues is apparent in the structure<sup>8</sup> and properties<sup>9</sup> of molecules that mimic collagen. To distinguish between the contributions of hydrogen bonding and inductive effects on collagen stability, we synthesized (ProFlpGly)<sub>10</sub>, where Flp is a 4(R)-fluoroproline residue<sup>8,9</sup>. We chose Flp because fluorine is the most electronegative element but does not form hydrogen bonds<sup>7</sup>.

The relative molecular mass of a triple helix of (ProFlpGly)<sub>10</sub> chains, as determined by sedimentation equilibrium in an analytical ultracentrifuge, is  $(8.0 \pm 0.1)K$ . In 50 mM acetic acid, this triple helix does not bind to 1-anilino-naphthalene-8-sulphonate, suggesting that its tertiary structure is packed tightly rather than

being like that of a molten globule<sup>10</sup>.

Most significantly, the circular dichroism (CD) spectrum of the (ProFlpGly)<sub>10</sub> triple helix is indistinguishable from that composed of (ProHypGly)<sub>10</sub> or (ProProGly)<sub>10</sub> chains (Fig. 1). These spectra are diagnostic of a collagen triple helix<sup>11</sup>.

Flp residues enhance triple-helix stability. The values of  $T_m$  for the three triple helices differ dramatically, increasing in the order (ProProGly)<sub>10</sub> < (ProHypGly)<sub>10</sub> < (ProFlpGly)<sub>10</sub> (Fig. 2). The thermal stability of (ProFlpGly)<sub>10</sub> triple helices far exceeds that of any known collagen of similar size.

We conclude that water bridges do not contribute significantly to collagen stability. Rather, collagen stability relies on previously unappreciated inductive effects. These inductive effects probably enhance collagen stability by favouring the requisite *trans* conformation of the hydroxyprolyl peptide bond<sup>9</sup>.

Because a fluorine atom exerts a greater inductive effect than does a hydroxyl group, Flp residues contribute more to collagen stability than do Hyp residues. Accordingly, chemical modification of the Hyp hydroxyl group with an electron-withdrawing substituent could enhance the stability of natural collagen, allowing new biomaterials to be produced<sup>12</sup>.

Steven K. Holmgren, Kimberly M. Taylor,  
Lynn E. Bretscher, Ronald T. Raines

Departments of Biochemistry and Chemistry,  
University of Wisconsin–Madison, Madison,  
Wisconsin 53706, USA

e-mail: raines@biochem.wisc.edu

1. Nimmi, M. E. (ed.) *Collagen* 1–4 (CRC, Boca Raton, Florida, 1988).
2. Brodsky, B. & Shah, N. K. *FASEB J.* **9**, 1537–1546 (1995).
3. Berg, R. A. & Prockop, D. J. *Biochem. Biophys. Res. Commun.* **52**, 115–120 (1973).
4. Bella, J., Brodsky, B. & Berman, H. M. *Structure* **3**, 893–906 (1995).
5. Bella, J., Eaton, M., Brodsky, B. & Berman, H. M. *Science* **266**, 75–81 (1994).
6. Engel, J., Chen, H.-T., Prockop, D. J. & Klump, H. *Biopolymers* **16**, 601–622 (1977).
7. Ojima, I., McCarthy, J. R. & Welch, J. T. (eds.) *Biomedical Frontiers of Fluorine Chemistry* (Amer. Chem. Soc.,



- Washington, DC, 1996).
8. Panasiuk, N., Jr, Eberhardt, E. S., Edison, A. S., Powell, D. R. & Raines, R. T. *Int. J. Pept. Protein Res.* **44**, 262–269 (1994).
  9. Eberhardt, E. S., Panasiuk, N., Jr & Raines, R. T. *J. Am. Chem. Soc.* **118**, 12261–12266 (1996).
  10. Semisotnov, G. V. *et al. Biopolymers* **31**, 119–128 (1991).
  11. Shah, N. K., Ramshaw, J. A. M., Kirkpatrick, A., Shah, C. & Brodsky, B. *Biochemistry* **35**, 10262–10268 (1996).
  12. Werkmeister, J. A. & Ramshaw, J. A. M. (eds) *Collagen Biomaterials* (Elsevier Science, Barking, Essex, 1992).

## Gene translocation links insects and crustaceans

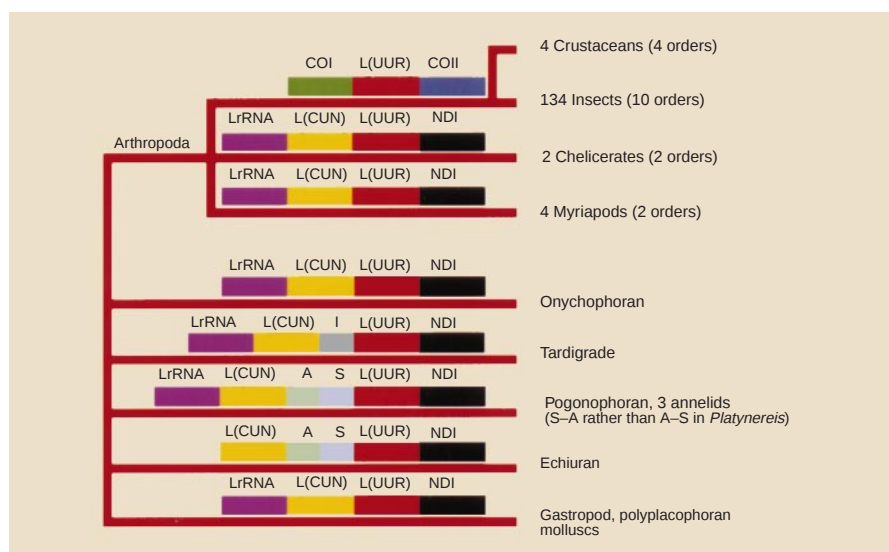
The evolutionary relationships among the four major lineages of arthropods remain controversial, despite extensive study. We report here a derived gene rearrangement common to insects and crustaceans but absent in the other arthropod groups. This finding strongly supports an insect–crustacean evolutionary lineage that is separate from those leading to myriapods and chelicerates.

The four major arthropod groups are Chelicerata (such as scorpions and horseshoe crabs), Crustacea (such as crabs and brine shrimp), Myriapoda (such as centipedes and millipedes), and Insecta (such as flies and beetles). Much of arthropod evolution remains contentious but, until recently, there has been general agreement that myriapods are the closest relatives of insects, forming a group known as the Atelocerata.

However, several recent morphological and molecular comparisons suggest that crustaceans, rather than myriapods, are the sister group to insects. If this is the case, some characteristics shared by insects and myriapods (such as a tracheal system for respiration, Malpighian tubules for excretion, and unbranched legs) then become examples of convergent evolution, perhaps as adaptations to life on land. Similar features are found among terrestrial chelicerates, where their occurrence is already viewed as convergent.

In a previous study<sup>1</sup>, we reported that insects and crustaceans share a derived location for the gene encoding mitochondrial leucine transfer RNA, designated *L(UUR)*, as compared with its primitive location in a chelicerate, an onychophoran, and several non-arthropod metazoans. However, in that study we were unable to associate a myriapod (*Thyropygus*) with either group.

Now, further mitochondrial DNA (mtDNA) sequence for *Thyropygus* and for three other myriapod species allows us to make that association. These four myriapods share the gene arrangement *LrRNA-L(CUN)-L(UUR)-ND1*, which is almost certainly primitive<sup>1</sup>. Our earlier misinterpretation of the *Thyropygus* sequence was due to a similarity of the amino-acid sequence



**Figure 1** Relative location of the *L(UUR)* gene for 153 taxa. The primitive location is identified in many non-arthropods and is retained in the mitochondrial genomes of the chelicerates and myriapods. An insect–crustacean clade is identified by the shared translocation of *L(UUR)* to the position between *COI* and *COII*. For 49 of these insect taxa, only the gene arrangement *L(UUR)-COII* has actually been determined<sup>9</sup>. For the chelicerates, myriapods, onychophoran, tardigrade, and echiuran, *COI-COII* are directly adjacent without any intervening tRNA genes; other taxa have an unrelated tRNA here. The three crustaceans and ten insects share the gene arrangement *LrRNA-L(CUN)-ND1*. For all arthropods for which the relative locations of these two gene blocks have been determined<sup>2–7</sup>, they are separated by more than 2.5 kilobases and are encoded on opposite DNA strands. Data are from published sources for three crustaceans (*Homarus*<sup>1</sup>, *Daphnia*<sup>1</sup>, *Artemia*<sup>2</sup>), the insects<sup>8</sup> (additional citations available from JLB), a chelicerate (*Limulus*<sup>2</sup>), onychophoran (*Euperipatoides*<sup>3</sup>), annelid (*Lumbricus*<sup>3</sup>), gastropod (*Plicopurpura*<sup>3</sup>) and polyplacophoran (*Katharina*<sup>3</sup>). Sequences determined here are for a remipede crustacean (*Speleonectes*), chelicerate (*Pandinus*), four myriapods (*Thyropygus*, *Lithobius*, *Spirostrephon*, *Narceus*), tardigrade (*Thulinia*), pogonophoran (*Galathea-olinum*), two annelids (*Helobdella*, *Platynereis*) and echiuran (*Urechis*). Sequences were determined from DNA fragments amplified using the polymerase chain reaction with primers made to conserved gene regions. Mitochondrial DNA typically contains 37 genes, only a subset of which is shown here (gene abbreviations are as published<sup>9</sup>).

inferred from the terminal 39 nucleotides of the *L(UUR)* gene. These nucleotides are in frame with the actual start site of the *ND1* gene and match 5/13 of the corresponding residues of the *ND1* gene in *Drosophila*.

The most parsimonious explanation for the gene arrangement data (Fig. 1) is that a single translocation of the *L(UUR)* gene occurred in a common lineage that led, after it split from the other lineages shown, to crustaceans and insects. This signature of common evolutionary history persists in the mtDNAs of these groups today.

If myriapods and insects were sister groups, either this tRNA translocation would need to have occurred twice identically in the lineages leading to insects and crustaceans, or it would need to have reverted to its primitive state in the myriapods. Each of these explanations would require an identical complex process.

Furthermore, this gene has translocated to a position remote from the original one. So the process is not a simple exchange of positions between neighbouring genes, nor does it involve genes adjacent to a large non-coding region, either of which might increase the frequency of gene rearrangement.

Another argument against the change being convergent is the infrequency of rearrangements among arthropod mitochondrial DNAs. Complete arrangements of all 37 mitochondrial genes have been determined for six arthropod genera: one chelicerate (*Limulus*<sup>2</sup>); one crustacean (*Artemia*<sup>2</sup>); and four insect (*Drosophila*<sup>4</sup>, *Locusta*<sup>5</sup>, *Anopheles*<sup>6</sup>, *Apis*<sup>7</sup>). The *Drosophila* arrangement differs from that of *Limulus* only in the location of *L(UUR)*, from that of *Artemia* only in the location of the tRNA gene block *I-Q*, and from those of *Locusta*, *Anopheles*, and *Apis* by one, two and eight tRNA translocations, respectively. There is no evidence of 'hot spots' for tRNA gene translocations in these genera, although the sample size is small.

Rearrangements of the 37 genes typical of metazoan mtDNA appear to be unique, rare events, unlikely to be duplicated by convergence, stable once they have occurred, and easily recognized because of the homology of mitochondrial genes across the Metazoa<sup>1</sup>. Our phylogenetic interpretation requires no convergence in any taxon for which data are available (more than 200 taxa representing 8 phyla).

We believe that this synapomorphy

strongly supports an insect–crustacean clade that excludes myriapods. Also, we anticipate that further study of the relative arrangements of the genes in metazoan mtDNA will help to clarify many other higher-level evolutionary relationships.

Jeffrey L. Boore, Dennis V. Lavrov,  
Wesley M. Brown

Department of Biology, University of Michigan,  
830 North University Avenue, Ann Arbor,  
Michigan 48109, USA  
e-mail: jboore@umich.edu

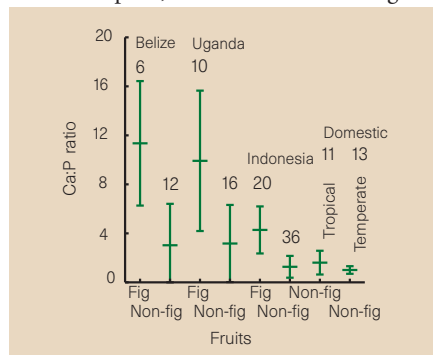
- Boore, J. L., Collins, T. M., Stanton, D., Daehler, L. L. & Brown, W. M. *Nature* **376**, 163–165 (1995).
- Stanton, J. L., Daehler, L. L. & Brown, W. M. *Mol. Biol. Evol.* **14**, 867–874 (1997).
- Valverde, J., Batuecas, B., Moratilla, C., Marco, R. & Garesse, R. *J. Mol. Evol.* **39**, 400–408 (1994).
- Clary, D. O. & Wolstenholme, D. R. *J. Mol. Evol.* **22**, 252–271 (1985).
- Flook, P., Rowell, C. H. F. & Gellissen, G. *J. Mol. Evol.* **41**, 928–941 (1995).
- Beard, C. B., Hamm, D. M. & Collins, F. H. *Insect Mol. Biol.* **2**, 103–124 (1993).
- Crozier, R. H. & Crozier, Y. C. *Genetics* **133**, 97–117 (1993).
- Boore, J. L. in *Current Topics on Molecular Evolution* (eds Nei, M. & Takahata, N.) 69–78 (Institute of Molecular Evolutionary Genetics, Pennsylvania State Univ., State College, PA, 1996).
- Boore, J. L. & Brown, W. M. *Genetics* **141**, 305–319 (1995).
- Boore, J. L. & Brown, W. M. *Genetics* **138**, 423–443 (1994).

## What's so special about figs?

Fruit-eating animals regularly prefer to eat figs even when other food is abundant. We propose that high calcium levels contribute to the desirability of figs as food for many forest animals.

There has been debate over the nutritional significance of the fig in the diet of frugivores<sup>1,2</sup> and over its importance as a 'keystone' species<sup>3–6</sup>. We compared the mineral composition of figs and other fruits from Belize, Indonesia and Uganda, and found that there was more than three times as much calcium in figs as in other fruits.

We analysed the mineral content of figs and other non-domestic fruit species from the neotropical, African and Asian regions



**Figure 1** Mean ( $\pm$  s.d.) calcium-to-phosphorus ratios for fig and wild non-fig fruits from Belize, Indonesia, and Uganda, and temperate and tropical domestic fruits. Numbers indicate species per sample.

**Table 1 Mineral analysis data**

| Ash                            | Ca   | Mg   | Na    | P    | Cu   | Fe   | Mn   | Zn   |
|--------------------------------|------|------|-------|------|------|------|------|------|
| <b>Belize fig</b>              |      |      |       |      |      |      |      |      |
| 8.13                           | 1.91 | 0.40 | 0.040 | 0.18 | 8.79 | 63.1 | 32.3 | 19.6 |
| 2.04                           | 0.70 | 0.14 | 0.016 | 0.05 | 2.29 | 18.8 | 38.2 | 4.3  |
| <b>Belize non-fig fruit</b>    |      |      |       |      |      |      |      |      |
| 5.99                           | 0.39 | 0.47 | 0.063 | 0.17 | 11.6 | 79.1 | 29.8 | 18.6 |
| 2.78                           | 0.35 | 0.29 | 0.037 | 0.12 | 6.38 | 55.6 | 30.4 | 10.4 |
| <b>Indonesia fig</b>           |      |      |       |      |      |      |      |      |
| 8.00                           | 1.21 | 0.25 | 0.060 | 0.33 | 10.0 | 65.7 | 10.9 | 26.1 |
| 3.08                           | 0.33 | 0.08 | 0.035 | 0.14 | 3.48 | 73.2 | 4.3  | 34.4 |
| <b>Indonesia non-fig fruit</b> |      |      |       |      |      |      |      |      |
| 6.77                           | 0.47 | 0.17 | 0.057 | 0.42 | 11.0 | 51.4 | 14.5 | 17.9 |
| 4.61                           | 0.37 | 0.10 | 0.049 | 0.35 | 7.54 | 42.0 | 14.9 | 11.8 |
| <b>Uganda fig</b>              |      |      |       |      |      |      |      |      |
| 9.40                           | 1.52 | 2.12 | 0.043 | 0.18 | 7.73 | 94.7 | 24.9 | 25.1 |
| 3.30                           | 0.55 | 5.48 | 0.048 | 0.08 | 2.44 | 68.6 | 13.5 | 18.2 |
| <b>Uganda non-fig fruit</b>    |      |      |       |      |      |      |      |      |
| 9.69                           | 0.48 | 0.36 | 0.014 | 0.14 | 8.43 | 122  | 42.1 | 22.2 |
| 6.69                           | 0.53 | 0.53 | 0.010 | 0.08 | 3.63 | 71.7 | 57.3 | 16.4 |

Values for ash, Ca, Mg, P and Na are expressed on a % dry-matter basis. Cu, Fe, Mn and Zn expressed as mg per g dry matter. Mean values are shown, with standard deviations underneath. Sample sizes are given in Fig. 1. Species information is available from the authors.

(Table 1), and 11 tropical and 13 temperate domestic fruit species. Regional differences were evident in the calcium, magnesium, iron, manganese, phosphorus and sodium content of all the fruits. This may have reflected varying soil fertility between collection sites.

However, between fig and non-fig fruits, differences in mineral concentrations were restricted to calcium. On average, figs contained calcium levels 3.2 times higher than other fruits — levels high enough to promote eggshell deposition in birds, and bone growth in birds and mammals<sup>7,8</sup>. Also, the ratio of calcium to phosphorus (a measure of calcium availability<sup>7,9</sup>) was 3.7 times higher in figs than in other fruits (Fig. 1). Figs from Sulawesi, Indonesia contained more calcium relative to the calcium availability in the soils, indicating that fig trees may selectively absorb calcium or allocate calcium to fruits.

Previous studies have shown that the protein, carbohydrate and lipid content of figs<sup>1,2</sup> are variable and not exceptionally high. Our study indicates that calcium concentration relative to phosphorus may be an important criterion for selection of the fruit.

Growth processes and egg-laying are accompanied by a rise in requirements for calcium and phosphorus<sup>7,8</sup> to lay down eggshell, aid metabolism, construct nucleic acids and form bone. Because most non-fig fruits, as well as seeds and invertebrates, are poor sources of calcium, many birds and mammals rely on calcium supplements such as mollusc shells, bone or soil to ensure adequate dietary calcium<sup>7,10,11</sup>. Others consume large quantities of figs throughout the year<sup>5,6,12</sup>. The biological availability of calcium in figs has not been determined, but animals whose diets are rich in figs are unlikely to suffer from calcium deficiency.

Terborgh<sup>3</sup> and others<sup>5,6,12</sup> suggest that figs constitute a 'keystone' plant resource for fruit-eating birds and mammals throughout the tropics. Figs display inter- and intraspe-

cific asynchrony in fruiting, tend to produce large crops, and show low interannual variation in fruit production<sup>13</sup>. These fruiting patterns make figs a reliable food source during times of general fruit scarcity.

Our results indicate that figs may be important throughout the year for maintaining an adequate balance of calcium among fruit-eating animals, thus fulfilling the role of a keystone plant resource for many animal species. These findings also suggest that the concentrations of specific minerals represent an important nutritional criterion for evaluating dietary choices across taxonomic groups and pantropical ecosystems.

Timothy G. O'Brien, Margaret F. Kinnaird,  
Ellen S. Dierenfeld

Wildlife Conservation Society,  
Bronx, New York 10460, USA  
e-mail: WCS-IP@indo.net.id

Nancy L. Conklin-Brittain,  
Richard W. Wrangham

Anthropology Department, Harvard University,  
Cambridge, Massachusetts 02138, USA

Scott C. Silver

Department of Biological Sciences,  
Fordham University, Bronx, New York 10460, USA

- Concklin, N. L. & Wrangham, R. W. *Biochem. Syst. Ecol.* **22**, 137–151 (1994).
- Bronstein, J. L. & Hoffman, K. *Oikos* **49**, 261–268 (1987).
- Terborgh, J. *Conservation Biology: The Science of Scarcity and Diversity* (ed. Soule, M.) 330–334 (Sinauer, Sunderland, MA, 1986).
- Gautier, A. & Michaloud, G. *Ecology*, **70**, 1826–1833 (1989).
- Lambert, F. & Marshall, G. J. *Ecology* **79**, 793–809 (1991).
- Kalko, E. J. *Biogeogr.*, **23**, 565–576 (1996).
- Robbins, C. T. *Wildlife Feeding and Nutrition* (Academic, New York, 1993).
- National Research Council. *Nutrient Requirements of Laboratory Animals*, Publ. No. 2767 (National Academy of Sciences, Washington DC, 1978).
- National Research Council. *Mineral Tolerance of Domestic Animals* (National Academy of Sciences, Washington DC, 1980).
- Graveland, J. & van der Wal, R. *Oecologia* **105**, 351–360 (1996).
- Barclay, R. M. R. *Symp. Zool. Soc. Lond.* **67**, 245–258 (1995).
- Kinnaird, M. F., O'Brien, T. G. & Suryadi, S. *Auk* **113**, 431–440 (1996).
- Janzen, D. H. *Annu. Rev. Ecol. Syst.* **10**, 13–52 (1979).



# Bright light comes around again

## Edmond Halley: Charting the Heavens and the Seas

by Alan Cook

Oxford University Press: 1998. Pp. 540.

£29.50, \$37.50

Owen Gingerich

"Wherefore if according to what we have already said [the comet] should return again about the year 1758, candid posterity will not refuse to acknowledge that this was first discovered by an Englishman." Candid posterity did not forget, and today Edmond Halley's comet carries more fame and mystique than any other.

Halley's brilliant prediction was in part a matter of luck: there is but one periodic comet visible with the naked eye, and Halley identified it after noticing the similarity of the orbital geometry of the comets of 1532, 1618 and 1682. But his success also came in part through following an avenue opened by the most important scientific book of the seventeenth century, Sir Isaac Newton's *Principia*. For historians of science, Halley's critical role as midwife in bringing out Newton's classic work looms even larger than his memorable but lesser work on the orbits of comets.

In the seventeenth century, however, Halley achieved international fame as a mathematician, although the importance of his geometrical publications has scarcely stood the test of time. More enduring was his brief observational campaign on St Helena in 1677–78, which resulted in the first published star catalogue made using telescopic sights, and his effort in editing *Philosophical Transactions*, which helped to hold together the Royal Society during otherwise precarious years. For his English contemporaries, he was particularly well known as a scientific seaman who charted currents, tides and magnetic bearings in two expeditions where he served as commander of the pink *Paramore*. And, as Sir Alan Cook now brings to light in this well-researched biography, Halley was appreciated in the circles of government as a diplomatic and maritime spy on the Dalmatian coast in the years 1702–03.

In the years that followed his maritime exploits, Halley was appointed professor of astronomy at the University of Oxford, and as henchman for Newton he became involved in an unseemly quarrel about publications with John Flamsteed, the first Astronomer Royal. On Flamsteed's death, Halley became the second Astronomer Royal and, at an age when most men would have long since retired, he followed the Moon through an entire 18-year nodal cycle.

Previous works about Halley, notably Eugene Fairfield MacPike's *Correspondence and Papers of Edmond Halley* (1932), assembled

some basic materials for a Halley memoir, but Cook has been far more diligent in his search through archives and record offices for Halley connections of all sorts. One of his most interesting findings suggests that the mysterious death of Halley's father in 1684 was not suicide, as claimed, but murder because he knew too much about the events in the Tower of London on the day the Earl of Essex died.

Cook's research also encompass virtually everything said about the astronomer's religious views, and he considers, but dismisses as biased, Flamsteed's view of Halley's smoking, "swearing like a sailor" and harbouring atheistical ideas. However, material to confirm Halley's orthodoxy is thin and circumstantial. Even more elusive is any insight into Halley's domestic life: virtually nothing is known of Mary Tooke, his wife of 53 years and the mother of his son and two daughters.

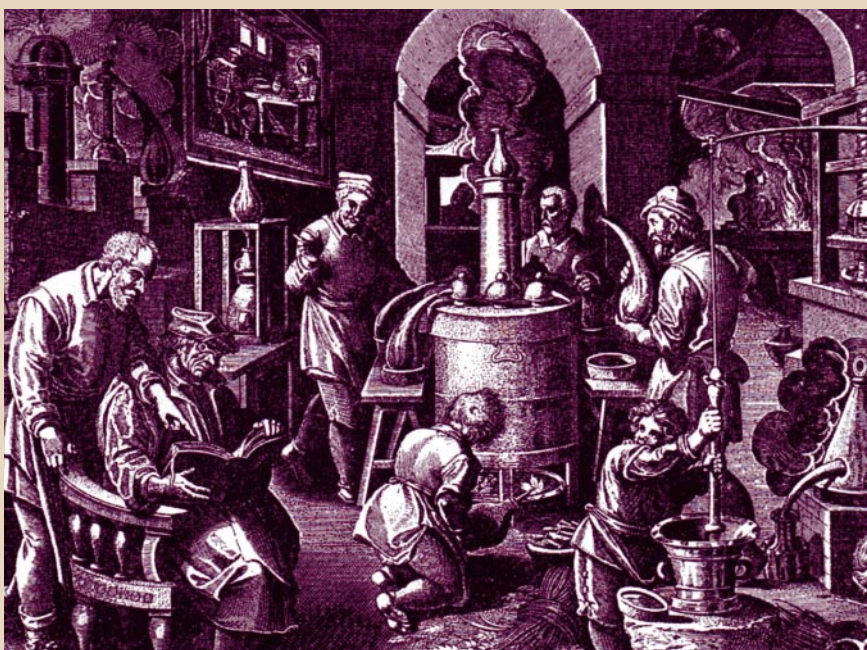
Cook's investigations of Parisian collections have helped to fill out the picture of Halley's trip in 1679 to Danzig to test and challenge the observational procedures of Johannes Hevelius, who used the naked eye in contrast to the telescopic sights used by Halley. Hevelius, a highly experienced 68-year-old observer, greatly impressed young Halley by the consistency of his measured positions. Archival research in Venice turned up fresh evidence for Halley's skill as a

cartographer of Mediterranean ports.

Occasional pieces of 'Halleyiana' have nevertheless escaped Cook's tight net. In 1684, from his private observatory in Islington (London), Halley made the first daytime telescopic observations of the planet Mercury, thereby enormously increasing the database for this elusive planet, an accomplishment that goes unrecorded here. Also, the helpful appendix "Correspondence of Halley not listed by MacPike" omits a brief but interesting 1695 letter about printing the *Synopsis of Comets* that has been published in the *Journal for the History of Astronomy* (16, 223–224; 1985).

There are also a few minor glitches, such as the names Andrew Sharp (instead of Abraham Sharp) and Christfried Kirck (instead of Kirch), and the incorrect statement that Hevelius had observed a transit of Venus. Newton's charming little drawing of the comet of 1680 exists in manuscript, but not in the *Principia*. Some of the transcripts would have been more readable if, for example, instead of the inscrutable "Lopps" the text had read "Lo[rds]hips", and US readers could have done with compact explanations of footpads and nonjurors.

Nonetheless, the overall impression created by the book is one of careful craftsmanship and brilliant stage setting to bring to life seventeenth-century England and the



## A man for all seasons

Although Newton is today chiefly remembered for his theoretical work *Principia*, he also spent much of his time doing experiments, sometimes using equipment like that shown in this artist's

impression of an alchemist's laboratory — or so argues Michael White in *Isaac Newton: The Last Sorcerer* (Fourth Estate, Addison-Wesley, £18.99, \$27).



London scene in particular. This serious and informative account promises to be the standard biography for many years to come. □

Owen Gingerich is at the Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts 02138, USA.

## Brains trust

### The Central Nervous System of Vertebrates

by R. Nieuwenhuys, H. J. ten Donkelaar and C. Nicholson

Springer: 1997. Three volumes. Pp. 2,219. £942, \$1,595

Glenn Northcutt

At a time when many scientists debate what constitutes an LPU — least publishable unit — this work, consisting of more than 2,000 pages in three volumes, must at least meet the minimal requirements. Such monumental surveys of the variation in the central nervous systems of vertebrates have been attempted several times.

In 1920, Cornelius Ubbo Ariëns Kappers, the first director of the famous Dutch Central Institute for Brain Research in Amsterdam, produced a two-volume work entitled *Vergleichende Anatomie des Nervensystems* which would form the core of a later expansion in 1936 by G. Carl Huber and Elizabeth C. Crosby. This three-volume work, *The Comparative Anatomy of the Nervous System of Vertebrates*, became a tombstone for the field — surely one could not write three volumes on the nervous systems of vertebrates without everything being said. Nothing could have been further from the truth. These surveys were based on descriptive histology which is particularly refractory to the analysis of microstructure and the interconnections of various parts of the nervous system.

Although the Kappers, Huber and Crosby volumes continued to be used during the 1960s, much of the original work and illustrations were badly out of date, and the material was not set within a broader biological context. These inadequacies were exacerbated by a renaissance of new techniques that allowed pathways to be experimentally established, neurotransmitters and their enzymes to be visualized, the ultrastructure of neurons and their processes to be examined in far greater detail than previously possible, and the refinement of electrophysiological techniques. Renewed interest in the development of nervous systems and their genetic underpinnings has further intensified the need for a new survey.

The new synthesis by R. Nieuwenhuys and colleagues is stunning and represents a level of scholarship that is rarely achieved in any field. The authors clearly outline the purpose and plan of this work by dividing it into three main sections: general features of

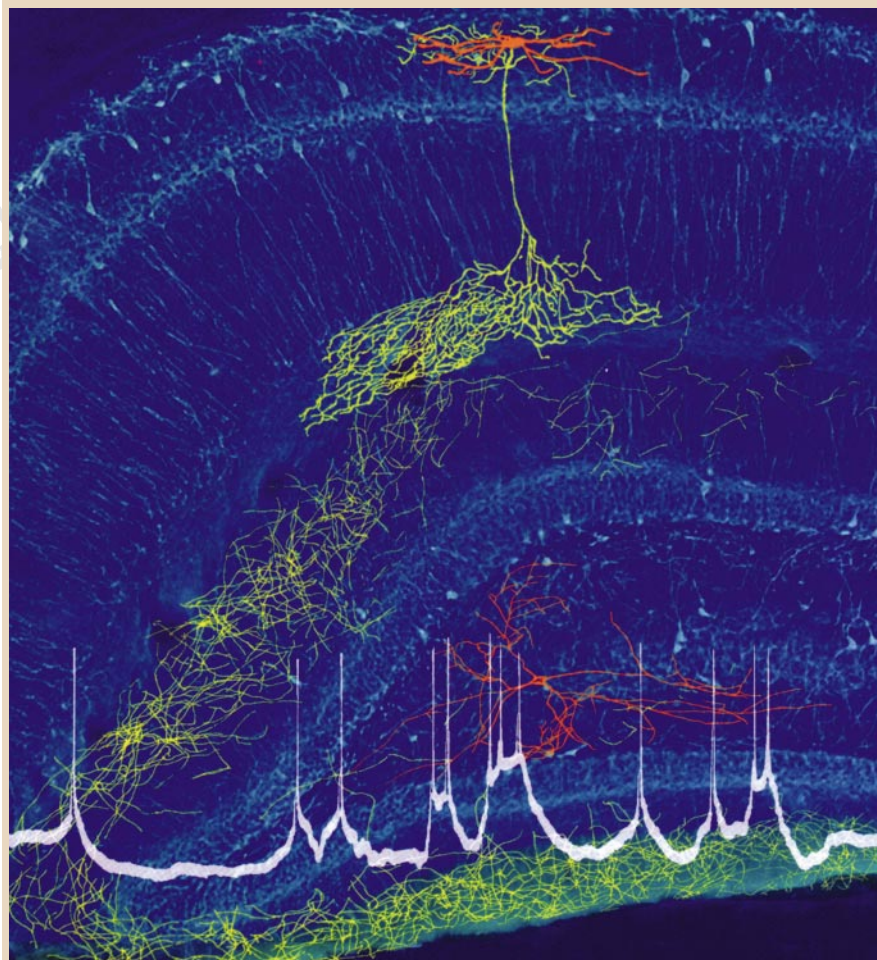
the cellular elements that form the nervous systems of vertebrates and the general principles that guide their interpretation; detailed descriptions of the brains of selected species of all groups of vertebrates; and possible directions for future research. The quality and uniformity of the illustrations are remarkable.

Seven chapters form the general introductory part of this work. The topics range from the structure and function of the cell types that compose nervous systems and their morphogenesis to the general biological framework within which comparative neurobiologists analyse these systems. These chapters will be invaluable to the novice, who may know little about the organization and functions of neural cells, as well as to more advanced students who understand the basic structure of neural cells but have little idea of how they are generated.

The final chapter in this section outlines the goals of comparative neurobiologists and discusses how a new method of comparison,

cladistics, initially generated in another biological field, is being used by neurobiologists to gain a deeper understanding of neural evolution. This chapter will be of immense value to contemporary neurobiologists and, ironically, will also constitute a comprehensive statement for historians of the biological sciences who could otherwise only guess the goals and programmes of comparative neurobiology at the end of the twentieth century.

The second section has 15 chapters summarizing the central nervous systems of amphioxus and representatives of all groups of vertebrates. These chapters follow a general pattern in which the life history of one or more species of a group is briefly summarized, and standard views of their brains are presented, followed by a sequence of drawings of transversely sectioned tissue from the spinal cords and brains of these same species. The authors have adopted a standard way of illustrating the various brain levels by primarily using line drawings rather than photomicrographs. Some readers will



## Looking closely at neuroscience

This section through the immunostained hippocampus of a rat adorns the cover of *Foundations of Neurobiology* (W. H. Freeman, \$55.95, £32.95). This broad-based neurobiology textbook, written by Fred Delcomyn for

students early in their training, will prepare them for further study in neuroscience. It is clearly laid out, with key terms in bold and generous use of tables and illustrations, and contains a useful glossary.



object to this convention, because it represents one more step away from the actual material, but it can be justified because it provides uniformity and a high level of illustration that is almost always lacking, even in major anatomical works.

The last part of each chapter summarizes functional aspects and the phylogenetic significance of the brains examined. It is here that the authors' scholarship shines most brightly. They present a balanced and comprehensive survey of current knowledge of the particular group of vertebrates described, rather than simply presenting the results of their own research. In rare cases, however, the authors adopt such a neutral position on critical issues that an informed reader will be frustrated.

The third part of this work focuses on brain size in vertebrates, and attempts to recognize major patterns or principles of organization of the major brain divisions. The chapter on brain size beautifully summarizes a vast literature, and the authors rightfully conclude that no single biological factor is likely to account for the vast range of brain sizes among vertebrates. The final chapter, "The Meaning of it All", gives a bird's eye view of the vast range of brain variation that exists among living vertebrates. In so doing, it poses questions that should continue to excite future generations rather than capping the field with a second tombstone. □

Glenn Northcutt is in the Department of Neurosciences, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0201, USA.

## Living with the moose

The North American moose, *Alces alces*, is the largest living deer, and roams freely across a band of forest covering Newfoundland, the Rocky Mountains, Alaska and as far south as Colorado. If the one on the right seems a little wary, it is perhaps no surprise.

Prized by native Americans and an important part of their culture, the moose was hunted by European settlers almost to extinction by the start of this century; the population has since recovered to around a million. Hunting still goes on (but licensed now), and man kills a further 3,500 in road accidents each year.

Albert W. Franzmann and Charles C. Schwartz have brought together 21 wildlife researchers to describe the moose's biology and ecology and analyse the strategies of managing and living alongside the animal. The result is the comprehensive *Ecology and Management of the North American Moose* (Wildlife Management Institute / Smithsonian Institution Press, \$59.95, £46.75).



## Is meta better?

### How Science Takes Stock: The Story of Meta-Analysis

by Morton Hunt

Russell Sage Foundation: 1997. Pp. 210.

\$29.95, £24

Steve Blinkhorn

A new kind of alchemy is abroad in the world. Arcane, esoteric and mesmerizing, it promises not to turn base metal into gold, nor to provide an elixir of youth, but to transmute statistical sows' ears into scientific silk purses. Meta-analysis, the systematic synthesis of otherwise inconclusive research findings into hard results, has, in the space of barely two decades, come from a standing start to challenge the blind clinical trial as the principal arbiter of truth and the touchstone of knowledge across a range of biomedical and social sciences.

Precisely because it is in these fields that hard, fast and exact findings are hard to come by, methodology meets morality with peculiar poignancy. Why go to the trouble and expense of blind experimentation, and subject a substantial proportion of a research

sample to suboptimal treatment, when a little fastidious dredging through inconclusive research literature can be turned, with a little statistical wizardry, into a clear result? The prospect of being able to complete a paper with the words "further research is not necessary" can be too tantalizing to ignore, particularly where conclusions affect the lives and prospects of real people.

The practice of meta-analysis is by no means a soft option. Done properly it requires long and tedious library research to identify studies that meet the criteria for statistical synthesis. In many respects the statistical procedures that form the visible element of meta-analysis are the least troublesome part of the process. The problem for researchers — and journal editors — is that the process of screening suitable results for inclusion in the synthesis is necessarily laborious and painstaking, not to say difficult to replicate. Journal referees can hardly be expected to undertake sufficient checking to ensure that the research literature has the status claimed by the meta-analyst. The audience is therefore more than usually reliant on the honesty, accuracy and competence of the research team in evaluating their claims.

An accessible and balanced account of the practice of meta-analysis may be thought overdue. Sadly, this book does not fit the bill. An openly partisan, not to say propagandist, introduction to the field, it has the same cringe-inducing quality found in certain excessively contemporary translations of the New Testament of the Bible.

Morton Hunt has provided subversive footnotes; subversive, that is, of his own purpose. For instance, in a dustjacket puff, Richard Light, one of the leading figures in meta-analysis, tells us that Hunt has "gotten all the important details right". But a footnote in the text tells us that in conventional hypothesis testing "p is the probability that [the] null hypothesis is true". If an understanding of the rudiments of inferential statistics is as unimportant to meta-analysis as this potted travesty of wisdom would suggest, we might as well abandon the pretence that undergraduates can get to grips with conventional hypothesis testing.

A little later in the book, the tale is told of how meta-analysis was first developed to rebut Hans Eysenck's claim that psychotherapy is ineffective. A footnote explains that the British use of the term 'psychotherapy'



excludes behavioural treatments, whereas the US usage includes them. It appears to have escaped Hunt's attention that Eysenck's argument was precisely that behavioural treatments succeed where 'talk therapies' fail.

It is not only in content that this account disappoints. The prose style owes more than a little to inspirational self-help journalism of a kind to be found in many a glossy monthly. In what, presumably, is intended as an attempt to humanize a dry subject, researchers prowl their offices like caged tigers, have merry twinkles in their eyes and have grandfatherly looks and gentle manners that give no hint of their international eminence. We are invited to believe that the strain of conducting a major meta-analysis together did not affect a couple's marriage, although they divorced a little while after it was finished for Completely Different Reasons. Political correctness has gained not even a toe-hold, for we read of slim, pert women who are 40 years old but look 30. Research is the more credible, perhaps, when conducted by the pert and neotenous.

Ultimately, the issue raised by meta-analysis concerns the kinds of question science should try to answer. At one extreme there are questions of the kind "Does psychotherapy work?", to which the most likely, if unhelpful, answer is "Probably sometimes, but by no means always". At the other extreme, "Is the mortality associated with lumpectomy greater than that associated with radical mastectomy?" admits a degree of precision. The former question suggests the need to clarify concepts and distinguish circumstances and methods, whereas the latter suggests that a point estimate of a probability is appropriate.

Unfortunately, even when questions are put precisely, the answers provided by meta-analysis do not always concur with the findings of large-scale conventional trials. In a recent review in *New England Journal of Medicine*, comparison with later randomized trials showed that following the results of meta-analyses would have led to ineffective treatment in 32% of cases, and the rejection of a useful treatment in 33% of cases.

In practice, for the foreseeable future, the safest approach is to treat meta-analysis as providing superior guidance in the formulation of hypotheses and the design of conclusive studies, rather than providing conclusive answers itself. It would be a shame to lose the extra leverage it can provide, but this account of its virtues, far from being convincing, serves only to undermine confidence in its usefulness. □

Steve Blinkhorn is at Psychometric Research and Development Ltd, Bawn House, The Maltings, St Albans, Hertfordshire AL1 3HT, UK.

## Lighting up the dark

### The Quantum Challenge

by George Greenstein and Arthur G. Zajonc  
*Jones and Bartlett: 1997. Pp. 204. \$50.*

Philippe Grangier

The conceptual foundations of quantum mechanics have been elaborated largely as a result of *Gedanken* experiments — thought experiments that are physically possible but are not designed to be carried out. Is there anything to learn by implementing a *Gedanken* experiment? Theoreticians may think not, but George Greenstein and Arthur G. Zajonc say that there is. To support this claim they have taken real experiments as a basis for presenting a detailed analysis of the mysteries of quantum mechanics.

The initial chapters of the book discuss experiments that are now more than ten years old, and the subjects are mature enough to be analysed with great clarity. They include wave-particle duality, illustrated through single-photon and delayed-choice interference experiments, and the uncertainty principle, associated with squeezed light. Special attention is devoted to entangled states and the Einstein-Podolsky-Rosen argument, accompanied by a convincing presentation of the experimental tests of Bell's inequalities.

The final chapters, dealing with measurement and decoherence and dissecting Schrödinger's cat, seem less polished, perhaps because direct experimental evidence was scarce when the book was written. But they nevertheless provide an interesting account of the slow evolution of experimental realization, and help one to appreciate the objectives and subtleties of the achievements that followed.

The book could be read with great profit by students who, after a term of a course on quantum mechanics, have grasped the basic calculatory techniques. A lot can be learnt about what practical quantum mechanics really is from the explanation and contemplation of elegant and subtle experiments.

The conceptual difficulties, carefully listed by the authors, have often been considered as "the dark side of quantum mechanics" and have been hidden away in standard textbooks. In that respect, the great intellectual honesty of the authors, who never hide the difficulties but also never overstate them, will be of great help to young readers. More knowledgeable readers will also learn a lot, but those looking for foggy philosophical arguments may be disappointed. The picture of quantum mechanics that emerges is far from obvious, but it is drawn from the weight of experimental reality. □

Philippe Grangier is at the Institut d'Optique, BP 147, F-91403 Orsay cedex, France.

## New in paperback

### The Origins of Virtue: Human Instincts and the Evolution of Cooperation

by Matt Ridley

Penguin, £8.99, \$13.95

"The book is extremely well written with the sort of anecdotal detail and with that make for lively reading even when the most abstract topics are being treated", Frans B. M. de Waal, *Nature* 383, 785 (1996).

### Himself and Other Animals: A Portrait of Gerald Durrell

by David Hughes

Pimlico, £7.99

Candid memoir of the famous conservationist by his friend and publisher of over 40 years.

### Unravelling DNA: The Most Important Molecule

by Maxim D. Frank-Kamenetskii

Addison-Wesley, \$15

Reviewed by David Lilley in *Nature* 367, 330 (1994).

### The Synaptic Organization of the Brain, 4th edn

edited by Gordon M. Shepherd

Oxford University Press, \$45

### A Man on the Moon: The Voyages of the Apollo Astronauts

by Andrew Chaiken

Penguin, \$15.95

### The History of Pain

by Roselyne Rey

Harvard University Press, \$18.95, £12.50

Cultural history. Reviewed by Lucy Bending in *Nature* 379, 129 (1995).

### The Evolutionary Synthesis: Perspectives on the Unification of Biology

edited by Ernst Mayr and William B. Provine

Harvard University Press, \$19.95, £13.50

Classic essays by key players in the 'modern synthesis' of evolutionary biology.

### Albert Einstein and the Frontiers of Physics

by Jeremy Bernstein

Oxford University Press, £7.99

Biography for beginners.

### Prisons of Light: Black Holes

by Kitty Ferguson

Cambridge University Press, £15.95, \$24.95

Popular introduction.

### The Cells of the Body: A History of Somatic Cell Genetics

by Henry Harris

Cold Spring Harbor Laboratory Press, \$29

# Microquasars in our Galaxy

I. F. Mirabel & L. F. Rodríguez

**Microquasars are stellar-mass black holes in our Galaxy that mimic, on a smaller scale, many of the phenomena seen in quasars. Their discovery opens the way for a new understanding of the connection between the accretion of matter onto black holes and the origin of the relativistic jets observed in remote quasars.**

Discovered more than 30 years ago, quasars remain some of the most mysterious objects in the Universe. It is widely believed that they are powered by black holes of several million solar masses or more that lie at the centres of remote galaxies. Their luminosities are much larger than ordinary galaxies like the Milky Way, yet originate from regions smaller than the size of the Solar System. Occasionally, quasars spout jets of gas that appear to move on the plane of the sky with velocities exceeding that of light (that is, with superluminal velocities). The extreme distance of quasars introduces many uncertainties into the interpretation of the source of energy and the nature of the ejecta that appear to be moving with superluminal speeds.

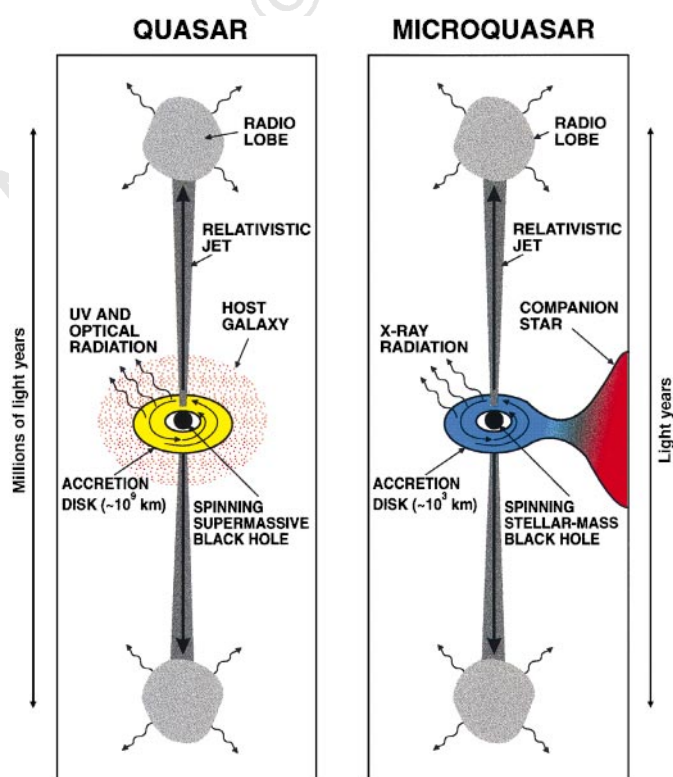
The recent finding in our own Galaxy of microquasars<sup>1-4</sup>, a class of objects that mimics—on scales millions of times smaller—the properties of quasars, has opened new perspectives for the astrophysics of black holes (see Fig. 1). These scaled-down versions of quasars are believed to be powered by spinning black holes<sup>5</sup> but with masses of up to a few tens times that of the Sun. The word microquasar was chosen to suggest that the analogy with quasars is more than morphological, and that there is an underlying unity in the physics of accreting black holes over an enormous range of scales, from stellar-mass black holes in binary stellar systems, to supermassive black holes at the centre of distant galaxies. As the characteristic times in the flow of matter onto a black hole are

proportional to its mass, variations with intervals of minutes in a microquasar correspond to analogous phenomena with durations of thousands of years in a quasar of  $10^9$  solar masses, which is much longer than a human lifetime. Therefore, variations with minutes of duration in microquasars could be sampling phenomena that we have not been able to study in quasars.

The repeated observation of two-sided moving jets in microquasars<sup>2,6</sup> has led to a much greater acceptance of the idea that the emission from quasar jets is associated with material moving at speeds close to that of light<sup>6</sup>. Furthermore, simultaneous multiwavelength observations of microquasars<sup>7,8</sup> are revealing the connection between the sudden disappearance of matter through the horizon of the black hole, with the ejection of expanding clouds of relativistic plasma.

## Superluminal sources

Superluminal motions have been observed in quasars for more than 20 years<sup>9</sup>. In the past, these motions had been used to argue that quasars cannot be as remote as believed, and that the use of redshifts and the Hubble expansion law to determine their distances was not fully justified. In the extragalactic case the jets are usually observed to be moving on only one side of the source of ejection, and it is not possible to know whether superluminal motions represent the propagation of waves through a slowly



**Figure 1** Diagram illustrating current ideas about quasars and microquasars (not to scale). As in quasars, the following three basic 'ingredients' are found in microquasars: (1) a spinning black hole, (2) an accretion disk heated by viscous dissipation, and (3) collimated jets of relativistic particles. But in microquasars the black hole is only a few solar masses instead of several million solar masses; the accretion disk has mean thermal temperatures of several million degrees instead of several thousand degrees; and the particles ejected at relativistic speeds can travel up to distances of a few light years only, instead of the several million light years as in some giant radio galaxies. In quasars matter can be drawn into the accretion disk from disrupted stars or from the interstellar medium of the host galaxy, whereas in microquasars the material is being drawn from the companion star in the binary system. In quasars the accretion disk has a size of  $\sim 10^9$  km and radiates mostly in the ultraviolet and optical wavelengths, whereas in microquasars the accretion disk has a size of  $\sim 10^3$  km and the bulk of the radiation leaves as X-rays.



### Box 1 Determining distances with special relativity

The superluminal motions can be fully understood in terms of a relativistic illusion in the observation of ejecta moving at close to the speed of light. The ejecta is moving so fast that it nearly catches up with its own radiation. After a time  $t$  from the moment of ejection, the plasma clouds, with a true velocity  $v$ , have moved a distance  $vt$ . As seen in projection by the observer, the displacement seems to be  $vt \sin \theta$ , where  $\theta$  is the angle between the line of sight and the axis of ejection (see Box Figure). However, because the approaching condensation is now closer to the observer by a distance  $vt \cos \theta$ , the time  $t'$  in which the observer sees the condensation move from the origin to its present position is smaller than  $t$  and is given by  $t' = t - (vt \cos \theta/c)$ . The apparent velocity of the approaching condensation is then  $v_a = v \sin \theta / (1 - (v \cos \theta/c))$ , that can exceed  $c$ . By a similar reasoning, the apparent velocity of the receding condensation is then  $v_r = v \sin \theta / (1 + (v \cos \theta/c))$ . If one can measure the proper motions in the sky of the approaching and receding ejecta,  $\mu_a$  and  $\mu_r$ , two independent equations are obtained:

$$\mu_a = \frac{v \sin \theta}{(1 - (v/c) \cos \theta) D}$$

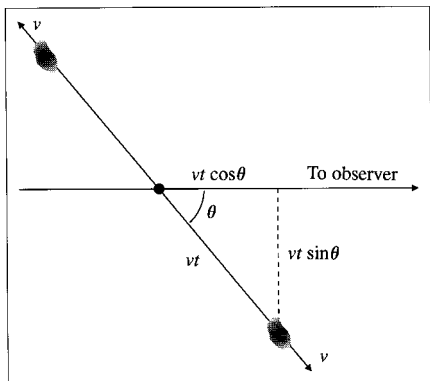
$$\mu_r = \frac{v \sin \theta}{(1 + (v/c) \cos \theta) D}$$

where  $\mu_a$  and  $\mu_r$  are in units of  $\text{rad s}^{-1}$ ,  $v$  is the true velocity of the ejecta,  $\theta$  is the angle between the line of sight and the ejection axis, and  $D$  is the distance to the source.

Measuring the wavelength  $\lambda_{a,r}$  of spectral lines (with rest wavelength  $\lambda_{\text{rest}}$ ) arising in either the approaching or receding jets, a third equation can be chosen from:

$$\frac{\lambda_{a,r}}{\lambda_{\text{rest}}} = \frac{1 \mp (v/c) \cos \theta}{[1 - (v/c)^2]^{1/2}}$$

One can then resolve the system of three equations and find the three unknowns:  $v$ ,  $\theta$  and  $D$ , the distance to the source. The physical configuration is shown in the Box Figure.



Additionally, for an object moving at relativistic speeds, the emitted radiation 'focuses' in the direction of motion (the so-called relativistic beaming), an effect that makes the approaching condensation look brighter than the receding condensation. In remote objects, such as quasars, the approaching ejecta can be observed more easily when their apparent brightness is very highly enhanced due to beaming to the observer. This Doppler-favouritism for the approaching jet implies the opposite effect for the receding jet, often rendering it undetectable in practice.

With the new technological capabilities in astronomy, this relativistic method to determine distances may be applied first to black-hole jet sources in binaries, and in the decades to come to quasars. The history of science shows that new methods to determine distances have been crucial in astronomical revolutions.

moving jet, or whether they reflect the actual motion of the sources of radiation.

In the context of the microquasar analogy, one then may ask if superluminal motions could be observed from black-hole binary systems known to lie in our own Galaxy. Among the handful of black holes of stellar mass known so far, the X-ray sources GRS1915+105 (ref. 10) and GRO J1655-40 (ref. 11) have indeed been identified at radio wavelengths as transient sources of superluminal jets<sup>2-4</sup>. The jets in the two sources move at 0.92 times the velocity of light and still hold the 'speed record' in the galaxy. GRO J1655-40 is known to be at a distance of 10,000 light years, and the apparent transverse motions of the ejecta in this source are the largest observed up to now from an object beyond the Solar System.

### Special relativity

The observation of superluminal double-sided jets in GRS1915+105 confirmed the existence of highly relativistic ejections of matter in the Universe and gave support to the microquasar analogy. Fig. 2 shows a significant ejection event observed from GRS1915+105 in March–April 1994. The brighter cloud (to the left) appears to move faster than the fainter cloud (to the right). This asymmetry in apparent motion and brightness can be explained in terms of relativistic aberration (see Box Figure). In all five main ejections observed so far from GRS1915+105, the clouds move with the same motions and approximately along the same direction on the sky<sup>6</sup>. The observation of counterjets in this microquasar allows a comparison of the parameters of the approaching and receding ejecta: this comparison shows that the emission arises in moving material and rules out the possibility of other explanations<sup>6</sup>, such as only rapid propagation of waves through a slowly moving jet.

As shown in Box 1, if the proper motions of the twin ejecta and the Doppler shift of spectral lines can be measured in the approaching and/or receding ejecta, the parameters of the system—in particular its distance—can be obtained<sup>12</sup>. The main observational difficulty resides in the detection of lines that are strongly Doppler broadened as they arise from plasma clouds that not only move but also expand at relativistic speeds.

### The central engine

Multiwavelength monitorings of the galactic superluminal sources have shown that hard X-ray emission is a necessary, but not sufficient, condition for the formation of collimated jets of synchrotron radio emission. In GRS1915+105, the relativistic ejection of pairs of plasma clouds have always been preceded by unusual activity in hard X-ray<sup>13</sup>. More specifically, the onset of the main ejection events seems to be related to the sudden drop from a luminous state in the hard X-ray<sup>14,15</sup> band. However, not all unusual activity and sudden drops in the hard X-ray flux appear to be associated with radio emission from relativistic jets. In fact, in GRO J1655-40 there have been several hard X-ray outbursts observed without subsequent radio flare/ejection events<sup>16</sup>.

What is the current belief as to how these relativistic jets are generated? Black holes, both in quasars and microquasars, are thought to be surrounded by an accretion disk. In the case of microquasars the disk is fed by magnetized gas drawn from a star that accompanies the black hole, forming a gravitationally bound binary system (Fig. 1). Finding a mechanism that creates a collimated flow that approaches the speed of light is a theoretical challenge faced more than two decades ago when extragalactic radio jets were first discovered. One possibility is magnetohydrodynamic acceleration in the accretion disk itself<sup>17</sup>. An alternative model is one that involves the enormous rotational energy of a spinning black hole<sup>18</sup>; the energy drawn from the black hole accelerates the magnetized plasma coming from the disk, and ultimately propels it into jets.

We next consider why some black-hole systems produce powerful jets while others apparently do not. The answer may reside in the

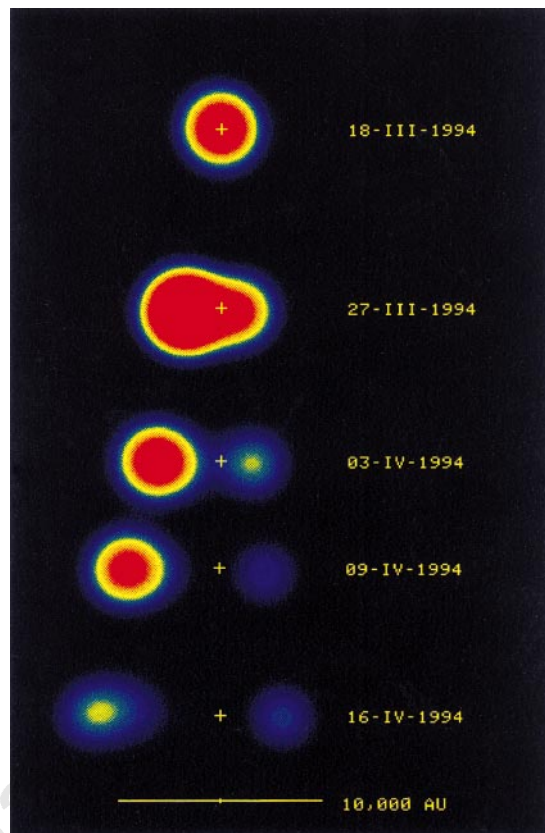
spin of the black hole. In a unification scheme<sup>19</sup>, black-hole sources of powerful collimated jets are rotating black holes with threading magnetic fields maintained by co-rotating accretion disks. Ejection of relativistic jets will take place preferentially in those sources where the spin of the black hole is close to its maximum value.

### Accretion onto the black hole

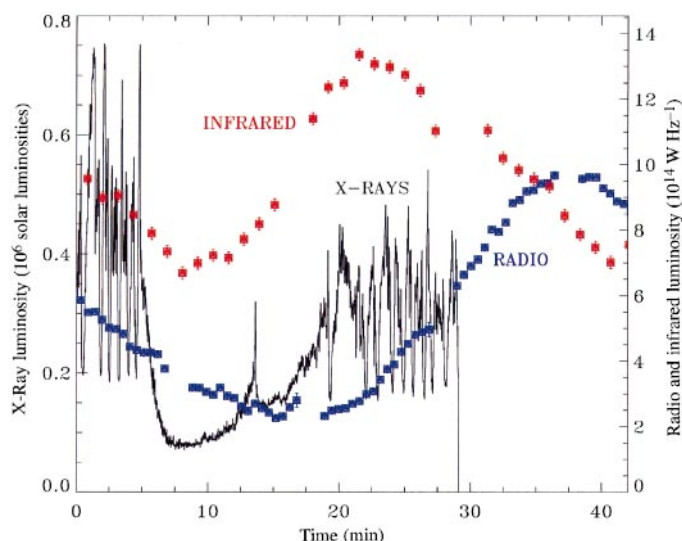
The X-ray power of the superluminal source GRS1915+105 shows a large variety of quasi-periodic oscillations (QPOs). Of particular interest is a class of oscillations with a maximum stable frequency of 67 Hz that has been observed many times, irrespective of the X-ray luminosity of the source<sup>20</sup>. It is believed that this fixed maximum frequency is a function of the fundamental properties of the black

hole, namely, its mass and spin. However, this particular frequency could be related to the last stable circular orbit around the black hole<sup>20</sup>, to a radial mode in general relativity disk seismology<sup>21</sup>, or to the relativistic dragging of the inertial frame around the fast-rotating collapsed object<sup>22</sup>. Theoretical work to distinguish between these alternatives will be important to estimate the spin of the black holes with masses that have been independently determined.

The episodes of large-amplitude X-ray flux variations on time-scales of seconds and minutes, and in particular, the abrupt dips observed in GRS1915+105 (see Fig. 3) are believed to be strong evidence for the presence of a black hole<sup>23–25</sup>. These variations could be explained if the inner part of the accretion disk goes temporarily into an advection-dominated mode<sup>26</sup>. In this mode, the time taken



**Figure 2** A pair of plasma clouds expelled from the microquasar GRS1915+105, the first superluminal source detected in the galaxy<sup>2</sup>. The observations were made at radio wavelengths, where red denotes the most intense emission. The cloud to the left appears to move away from the centre of ejection (yellow cross) at 125% the speed of light ( $1.25c$ ). An observer in the frame of one of these clouds would see the other receding at  $0.997c$ . It has been shown<sup>2</sup> that the asymmetries in velocity and brightness between the cloud to the left and the cloud to the right can be explained in terms of an antiparallel ejection of pairs of twin plasma clouds moving with bulk speeds of 92% that of light, at an angle of  $70^\circ$  with respect to the line of sight. These maps were obtained at intervals of  $\sim 1$  week with the Very Large Array of NRAO at a wavelength of 3.5 cm. At a distance of 40,000 light years from Earth, in 1 month the clouds moved apart in the plane of the sky a distance equivalent to 10,000 astronomical units (AU; 1 AU equals the mean radius of the Earth's orbit around the Sun).



**Figure 3** Oscillations in the luminosity at X-rays (2–60 keV), infrared ( $2\mu\text{m}$ ) and radio (6 cm) wavelengths of the galactic source of superluminal jets GRS1915+105 (ref. 7). The disappearance of the inner accretion disk (marked by the X-ray dip 7 min after the start of the set of observations shown in here), coincides with the beginning of the ejection of a relativistic plasma cloud (marked by the start of the infrared flare). As the ejected cloud expands it becomes transparent to radio waves, with a peak radio-wave flux that is delayed by 15 min relative to the infrared peak. The absence of X-ray data after 29 min is due to occultation of the source by the Earth.



for the energy transfer from ions (that get most of the energy from viscosity) to electrons (that are responsible for the radiation) is larger than the time of infall to the compact object. Then, the bulk of the energy produced by viscous dissipation in the disk is not radiated (as happens in standard disk models), but instead is stored in the gas as thermal energy. This gas, with large amounts of stored energy, is advected (transported) to the compact object. If the compact object is a black hole, the energy quietly disappears through the horizon, and one can observe sharp decays in the X-ray luminosity. In contrast, if the compact object is a neutron star, the thermal energy in the superheated gas is released as radiation when it collides with the surface of the neutron star and heats up. The cooling time of the neutron-star photosphere is relatively long, and in this case a slow decay in the X-ray flux is observed. Then, one would expect the luminosity of black-hole binaries to vary over a much larger range of values than that of neutron-star binaries for an interval of time of a few seconds.

### The formation of jets

During large-amplitude variations in the X-ray flux of GRS1915+105, remarkable flux variations on timescales of minutes have also been reported at radio<sup>27,28</sup> and near-infrared<sup>29</sup> wavelengths (see Fig. 3). These rapid flares are thought to come from expanding magnetized clouds of relativistic particles. Simultaneous observations in X-ray, infrared and radio wavelengths show that the ejection of relativistic clouds of plasma<sup>7,8</sup> takes place when the matter of the inner accretion disk suddenly disappears through the horizon of the black hole. Recent interferometric observations during episodic ejections with the Very Long Baseline Array by Dhawan and us, and with the British radio interferometer MERLIN by Fender *et al.* may show the evolution of the ejecta on timescales of hours and with spatial resolutions of tens of astronomical units.

Microquasars offer new opportunities to gain a general understanding of the relativistic jets seen elsewhere in the Universe. We expect that in the future they will also be used to investigate the physics of strong-field relativistic gravity near the event horizon of black holes. □

*I.F.M. is at DAPNIA/CEA Centre d'Études de Saclay, 91191 Gif/Yvette, France & Instituto de Astronomía y Física del Espacio, cc67, 1428 Buenos Aires, Argentina. L.F.R. is at the Instituto de Astronomía, UNAM, Morelia, Michoacán 58090, Mexico.*

1. Mirabel, I. F. *et al.* A double-sided radio jet from the compact Galactic Centre annihilator 1E1740.7-2942. *Nature* **350**, 215–218 (1992).

2. Mirabel, I. F. & Rodríguez, L. F. A superluminal source in the galaxy. *Nature* **371**, 46–48 (1994).
3. Tingay, S. J. *et al.* A relativistically expanding radio source associated with GRO J1655–40. *Nature* **374**, 141–143 (1995).
4. Hjellming, R. M. & Rupen, M. P. Episodic ejection of relativistic jets by the X-ray transient GRO J1655–40. *Nature* **375**, 464–467 (1995).
5. Shapiro, S. L. & Teukolsky, S. A. *Black Holes, White Dwarfs, and Neutron Stars* 441 (Wiley, New York, 1983).
6. Rodríguez, L. F. & Mirabel, I. F. Repeated relativistic ejections in GRS 1915+105. *Astrophys. J.* (submitted).
7. Mirabel, I. F. *et al.* Accretion instabilities and jet formation in GRS 1915+105. *Astron. Astrophys.* **330**, L9–L12 (1998).
8. Eikenberry, S. S. *et al.* Evidence for disk-jet interaction in the microquasar GRS 1915+105. *Astrophys. J.* **494**, L61–L64 (1998).
9. Pearson, T. J. & Zensus, J. A. in *Superluminal Radio Sources* (eds Zensus, J. A. & Pearson, T. J.) 1–11 (Cambridge Univ. Press, 1987).
10. Castro-Tirado, A. J. *et al.* Discovery and observations by Watch of the X-ray transient GRS 1915+105. *Astrophys. J. Suppl. Ser.* **92**, 469–472 (1994).
11. Zhang, S. N., Wilson, C. A., Harmon, B. A., Fishman, G. J. & Wilson, R. B. X-ray nova in Scorpius. *IAU Circ. No.* 6046 (1994).
12. Hjellming, R. M. & Johnston, K. J. An analysis of the proper motions of SS433 radio jets. *Astrophys. J.* **246**, L141–L145 (1981).
13. Harmon, B. A. *et al.* Hard X-ray signature of plasma ejection in the galactic jet source GRS 1915+105. *Astrophys. J.* **477**, L85–L90 (1997).
14. Foster, R. S. *et al.* Radio and X-ray variability of the galactic superluminal source GRS 1915+105. *Astrophys. J.* **467**, L81–L84 (1996).
15. Mirabel, I. F. *et al.* Infrared observations of an energetic outburst in GRS 1915+105. *Astrophys. J.* **472**, L111–L114 (1996).
16. Zhang, S. N. *et al.* in *Proceedings of the 4th Compton Symposium* (eds Dermer, C. D., Strickman, M. S. & Kurfess, J. D.) 141–162 (AIP Conf. Proc. 410, 1997).
17. Blandford, R. D. & Payne, D. G. Hydromagnetic flows from accretion discs and the production of radio jets. *Mon. Not. R. Astron. Soc.* **199**, 883–894 (1982).
18. Blandford, R. D. & Znajek, R. L. Electromagnetic extraction of energy from Kerr black holes. *Mon. Not. R. Astron. Soc.* **179**, 433–440 (1977).
19. Zhang, S. N., Cui, W. & Chen, W. Black hole spin in X-ray binaries: observational consequences. *Astrophys. J.* **482**, L155–L158 (1997).
20. Morgan, E. H., Remillard, R. A. & Greiner, J. RXTE observations of QPOs in the black hole candidate GRS 1915+105. *Astrophys. J.* **482**, 993–1010 (1997).
21. Nowak, M. A., Wagoner, R. V., Begelman, M. C. & Lehr, D. E. The 67 Hz feature in the black hole candidate GRS 1915+105 as a possible “diskoseismic” mode. *Astrophys. J.* **477**, L91–L94 (1997).
22. Cui, W., Zhang, S. N. & Chen, W. Evidence for frame-dragging around spinning black holes in X-ray binaries. *Astrophys. J.* **492**, L53–L57 (1998).
23. Greiner, J., Morgan, E. H. & Remillard, R. A. Rossi X-ray timing explorer observations of GRS 1915+105. *Astrophys. J.* **473**, L107–L110 (1996).
24. Belloni, T., Méndez, M., King, A. R., van der Klis, M. & van Paradijs, J. An unstable central disk in the superluminal black hole X-ray binary GRS 1915+105. *Astrophys. J.* **479**, L145–L148 (1997).
25. Chen, X., Swank, J. H. & Taam, R. E. Rapid bursts from GRS 1915+105 with RXTE. *Astrophys. J.* **477**, L41–L44 (1997).
26. Narayan, R., Garcia, M. R. & McClintock, J. E. Advection-dominated accretion and black hole event horizons. *Astrophys. J.* **478**, L79–L82 (1997).
27. Pooley, G. & Fender, R. P. Quasi-periodic variations in the radio emission from GRS 1915+105. *Mon. Not. R. Astron. Soc.* (in the press).
28. Rodríguez, L. F. & Mirabel, I. F. Sinusoidal oscillations in the radio flux of GRS 1915+105. *Astrophys. J.* **474**, L123–L125 (1997).
29. Fender, R. P., Pooley, G. G., Brocksopp, C. & Newell, S. J. Rapid infrared flares in GRS 1915+105: evidence for infrared synchrotron emission. *Mon. Not. R. Astron. Soc.* **179**, L65–L69 (1997).

**Acknowledgements.** We have benefited from discussions with colleagues at the Goddard workshop on Microquasars, and we thank J. Swank and A. Harmon for organizing this workshop.

Correspondence and requests for materials should be addressed to I.F.M. at Centre d'Études de Saclay (e-mail: mirabel@discovery.saclay.cea.fr).

# Electron transfer by domain movement in cytochrome *bc*<sub>1</sub>

Zhaolei Zhang<sup>\*†</sup>, Lishar Huang<sup>\*‡</sup>, Vladimir M. Shulmeister<sup>§</sup>, Young-In Chi<sup>‡</sup>, Kyeong Kyu Kim<sup>‡</sup>, Li-Wei Hung<sup>†</sup>, Antony R. Crofts<sup>||</sup>, Edward A. Berry<sup>‡</sup> & Sung-Hou Kim<sup>\*†‡</sup>

<sup>\*</sup> E. O. Lawrence Berkeley National Laboratory, <sup>†</sup> the Graduate Group of Biophysics, and <sup>‡</sup> Department of Chemistry University of California, Berkeley, California, 94720 USA

<sup>||</sup> Center for Biophysics and Computational Biology, University of Illinois at Urbana-Champaign, Illinois, 61801 USA

<sup>§</sup> This paper is dedicated to the memory of our friend and colleague Vladimir M. Shulmeister, a key member on the project until his untimely death on 27 September 1995.

**The cytochrome *bc*<sub>1</sub> is one of the three major respiratory enzyme complexes residing in the inner mitochondrial membrane. Cytochrome *bc*<sub>1</sub> transfers electrons from ubiquinol to cytochrome *c* and uses the energy thus released to form an electrochemical gradient across the inner membrane. Our X-ray crystal structures of the complex from chicken, cow and rabbit in both the presence and absence of inhibitors of quinone oxidation, reveal two different locations for the extrinsic domain of one component of the enzyme, an iron-sulphur protein. One location is close enough to the supposed quinol oxidation site to allow reduction of the Fe-S protein by ubiquinol. The other site is close enough to cytochrome *c*<sub>1</sub> to allow oxidation of the Fe-S protein by the cytochrome. As neither location will allow both reactions to proceed at a suitable rate, the reaction mechanism must involve movement of the extrinsic domain of the Fe-S component in order to shuttle electrons from ubiquinol to cytochrome *c*<sub>1</sub>. Such a mechanism has not previously been observed in redox protein complexes.**

Energy conversion in the biosphere occurs mainly through respiration and photosynthesis, and represents a flux several orders of magnitude greater than all anthropogenic energy usage. The underlying mechanism involves coupling electron transfer, along a chain of redox or photoredox enzymes, to proton translocation across an organellar membrane in which those redox components are embedded. This gives rise to a transmembrane electrochemical proton gradient, which can be coupled to energy-requiring processes, including synthesis of ATP—a principle first proposed by Mitchell in his chemiosmotic hypothesis<sup>1</sup>.

The central component of the electron-transfer chain in mitochondria and in many aerobic or photosynthetic bacteria is a complex of membrane proteins known as the cytochrome *bc*<sub>1</sub> complex, or ubiquinol:cytochrome *c* oxidoreductase (E.C. 1.10.2.2). This enzyme complex catalyses electron transfer from ubiquinol to a soluble cytochrome *c*; this transfer is coupled to translocation of two protons across the inner mitochondrial membrane per quinol oxidized<sup>2–4</sup>. The complex isolated from beef heart consists of 11 different polypeptides<sup>5,6</sup> and has a relative molecular mass (*M*<sub>r</sub>) of 240K (Table 1). There are four redox centres, namely, two haem groups, *b*<sub>H</sub> and *b*<sub>L</sub>, of cytochrome *b*, one haem group in cytochrome *c*<sub>1</sub>, and one iron-sulphur cluster of the Rieske protein. A mechanism accounting quantitatively for the proton translocation coupled to electron transport by this enzyme is a version of the 'proton-motive Q cycle' of Mitchell<sup>3,4</sup>. The mechanism also explains the pattern of inhibition by the ubiquinone analogues antimycin, stigmatellin, undecylhydroxydiazobenzothiazole, myxothiazol, and methoxyacrylo-stilbene, which bind specifically at one or the other of the two catalytic sites at which quinone is processed<sup>3,4</sup>.

Until recently, only a low-resolution structure for the cytochrome *bc*<sub>1</sub> complex from *Neurospora crassa* was available, from electron microscopy of two-dimensional crystals<sup>7</sup>. More recently, the *bc*<sub>1</sub> complex from beef mitochondria has been crystallized in three dimensions in a tetragonal space group<sup>8</sup> and in other space groups<sup>9,10</sup>. A partial structure of the complex has been reported from the tetragonal crystals<sup>11,12</sup>. In this structure, the extrinsic domain of the Rieske protein was too disordered to be traced, and cytochrome *c*<sub>1</sub> was only partially traced.

We have obtained other crystal forms<sup>13</sup> from other species, including one from chicken heart mitochondria that diffracts to

3.0 Å resolution. With these crystals we have now determined the structure of the complex, which includes the functionally important Rieske iron-sulphur protein and cytochrome *c*<sub>1</sub>. We were also able to assign three additional subunits (subunits 8, 10 and 11) that were not assigned before<sup>12</sup>.

A comparison of our structures in the presence and absence of various inhibitors shows that the extrinsic domain of the Rieske protein containing the iron-sulphur cluster assumes one of two conformations in the complexes. In one conformation, the iron-sulphur cluster is close to its electron acceptor, the haem group of cytochrome *c*<sub>1</sub>, but far from the presumed binding site of its electron donor, ubiquinol, in cytochrome *b*. In the other conformation, the iron-sulphur cluster is closer to cytochrome *b*, and farther from cytochrome *c*<sub>1</sub>. This conformation is similar to that found in the tetragonal beef crystals<sup>12</sup>.

We have located the binding sites for two Q<sub>o</sub>-site inhibitors, stigmatellin and myxothiazol, and for the Q<sub>i</sub>-site inhibitor, antimycin. The two Q<sub>o</sub>-site inhibitors bind in overlapping but not identical sites.

These two conformations for the iron-sulphur protein and three positions for binding of ubiquinone analogue inhibitors are compatible with all the reactions proposed by the Q-cycle mechanism

**Table 1 Subunits of the bovine heart cytochrome *bc*<sub>1</sub> complex**

| Subunit                              | Residues | <i>M</i> <sub>r</sub> |
|--------------------------------------|----------|-----------------------|
| 1 Core 1                             | 446      | 49,132                |
| 2 Core 2                             | 439      | 46,471                |
| 3 Cytochrome <i>b</i>                | 379      | 42,592                |
| 4 Cytochrome <i>c</i> <sub>1</sub>   | 241      | 27,288                |
| 5 Rieske Fe-S                        | 196      | 21,611                |
| 6 13.4K                              | 110      | 13,347                |
| 7 'Q-binding'                        | 81       | 9,590                 |
| 8 <i>c</i> <sub>1</sub> 'hinge'      | 78       | 9,170                 |
| 9 Fe-S presequence                   | 78       | 7,956                 |
| 10 <i>c</i> <sub>1</sub> -associated | 62       | 7,198                 |
| 11 6.4K                              | 56       | 6,363                 |
| Apo- <i>bc</i> <sub>1</sub>          | 2,166    | 240,718               |
| Fe <sub>2</sub> S <sub>2</sub>       |          | 76                    |
| Haem <i>c</i> <sub>1</sub>           |          | 616                   |
| Haem <i>b</i> <sub>H</sub>           |          | 616                   |
| Haem <i>b</i> <sub>L</sub>           |          | 616                   |
| Prosthetic groups                    |          | 2,014                 |
| Holo- <i>bc</i> <sub>1</sub> complex |          | 242,742               |



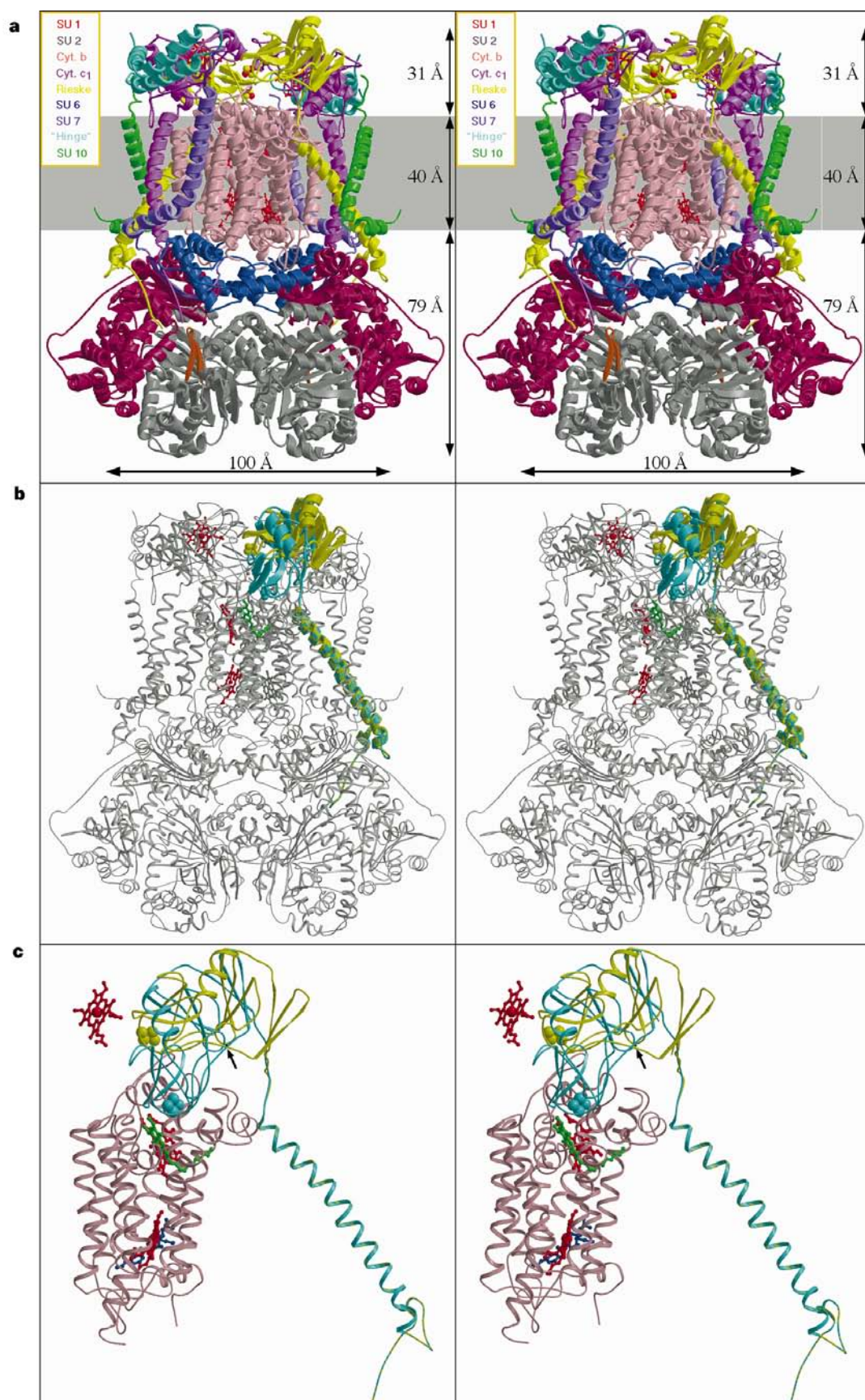
for electron transfer coupled to proton translocation. However, no one structure alone would be competent. We therefore propose that the reaction mechanism for electron transfer in the cytochrome  $bc_1$  complex requires a dramatic conformational change involving movement of the iron-sulphur protein.

A preliminary report of these results has appeared in the proceedings of a meeting<sup>14</sup>.

### Overall shape of cytochrome $bc_1$ complex dimer

In all crystals of the  $bc_1$  complex from three sources, the complex is

**Figure 1** Stereo-view ribbon diagrams of the  $bc_1$  complex. **a**, The native chicken  $bc_1$  dimer. The molecular two-fold axis runs vertically between the two monomers. The key for the colour code of each subunit is given in the inset. Quinones, phospholipids and detergent molecules are not shown for clarity. The presumed membrane bilayer is represented by a grey band. **b**, Two conformations of the Rieske protein in one monomer shown in the context of the entire dimer. One conformation found in our native chicken crystal (yellow) is superimposed on the other conformation (blue) from crystals grown in the presence of stigmatellin (green stick model). The haem groups (red) of cytochrome  $c_1$  and cytochrome  $b$  as well as the positions of the iron-sulphur cluster (orange and green balls) of the Rieske protein are shown. **c**, Isolated close-up view of the two conformations of the Rieske protein in contact with cytochrome  $b$  (pink), its associated haem groups (red), stigmatellin (green) and antimycin (purple). The isolated haem of cytochrome  $c_1$  (red, above) is also shown. SU, subunit; cyt, cytochrome.





present as a dimer (Fig. 1a), in which two monomers are related by a two-fold axis running vertically in the plane of the paper. The protein extends from the membrane 79 Å into the matrix space and 31 Å into the intermembrane region on either side of a transmembrane region 40 Å thick, giving a total length of 150 Å perpendicular to the membrane.

The overall shape of the dimer is similar to that described for the beef complex<sup>11,2</sup>, but we have modelled more protein in the intermembrane region. We have located subunits 1–8 and 10 in the electron density of the chicken crystal. We assign subunit 10 to the transmembrane helix labelled N1 in ref. 12, on the basis of good correlation between side chains in the chicken electron density and the beef subunit sequence. Subunit 11 seems not to be present in our preparation of the chicken enzyme, but is present in the beef and rabbit enzymes. It probably corresponds to the transmembrane helix labelled N2 in ref. 12, because this helix is present in the three crystal forms from the beef and rabbit enzymes and not in the chicken crystals.

Subunit 9 has not been assigned yet, and is also missing from the structure of ref. 12. This subunit is the presequence<sup>15</sup> of the Rieske protein, and is cleaved off by a matrix-located processing protease. We also see densities at several sites in the transmembrane portion. We attribute these densities to ubiquinone, detergents and phospholipids.

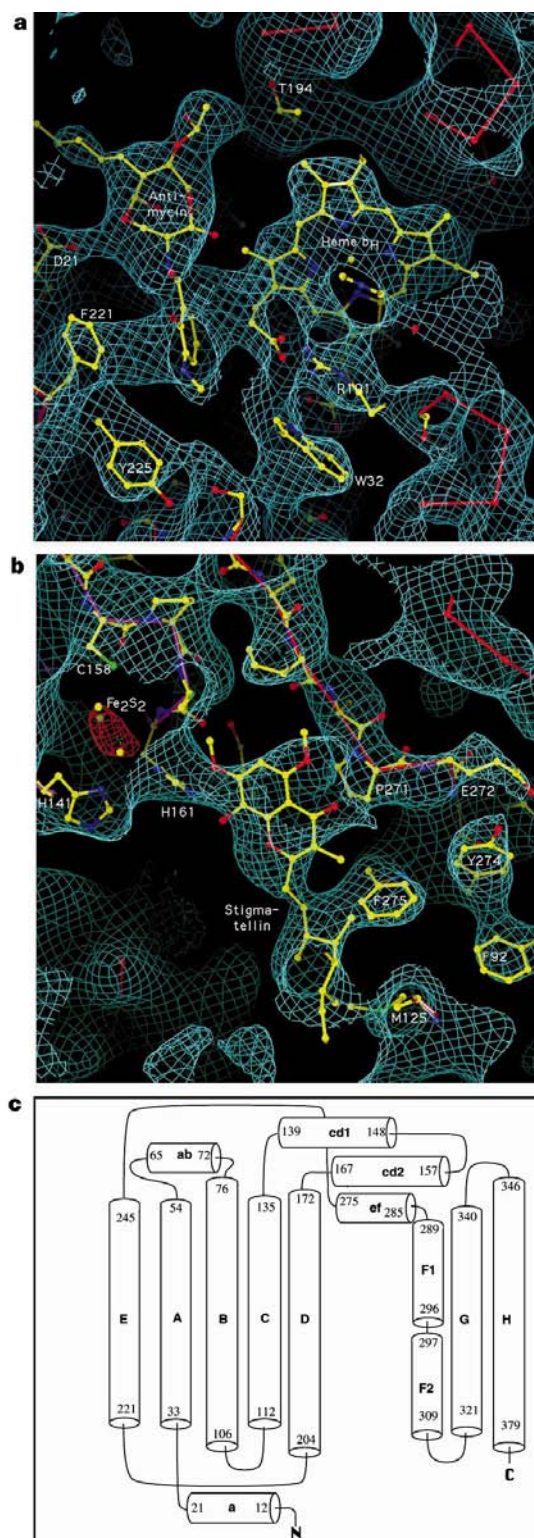
In the transmembrane domain the helices of the dimer fall into two clearly separated, packed bundles. We have arbitrarily divided the dimer so that one monomer corresponds to one packed bundle of helices in the transmembrane region.

### Inhibitor-binding sites

The presence of stoichiometric excesses of the inhibitors antimycin, myxothiazol, or stigmatellin during crystallization resulted in electron-density increases that could be interpreted as being due to the bound inhibitors. The general positions of the antimycin-, stigmatellin- and myxothiazol-binding sites are similar to those inferred from figures in refs 11, 12. Although the limited resolution does not allow the building of detailed atomic models, we have constructed speculative models consistent with the electron density. These inhibitor-binding sites, and especially the  $Q_o$  site, are the targets of drug-design efforts to produce environmentally safe and effective plant-protection fungicides for agriculture use<sup>16–18</sup>.

**The antimycin site.** On the basis of its mode of inhibition, antimycin is thought to bind at the  $Q_i$  site postulated in the Q-cycle mechanism. At this site, ubiquinone is reduced by electrons from cytochrome *b* accompanied by uptake of protons from the matrix space (resulting in proton translocation when ubiquinol is subsequently oxidized at the  $Q_o$  site, with proton release to the external medium). The antimycin-binding site (Fig. 2a) is near the high-potential haem group of cytochrome *b* ( $b_H$ ), in a cavity surrounded by the haem, the transmembrane helices A, D and E, and the amphipathic surface helix a (secondary structure is defined in Fig. 2c). There may be protonic connection to the matrix phase through or around conserved histidine 202. The close approach of the aromatic ring of the inhibitor to the haem group was expected from the effect of antimycin on the alpha absorption peak of  $b_H$  and the fluorescence quenching of antimycin when specifically bound at this site<sup>19</sup>. Residues F221 and T194 are also close enough to contact the inhibitor. One of the haem propionates is in van der Waals contact with the inhibitor and curves around to form an ion pair with R101. The conformation of this propionate, which differs from that depicted for the tetragonal beef crystals<sup>12</sup>, is the same in the absence of antimycin.

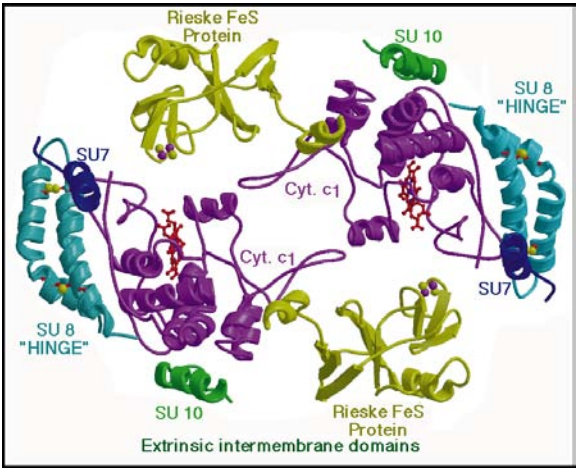
**The stigmatellin site.** The stigmatellin-binding pocket (Fig. 2b) is formed by the carboxy-terminal end of helix C, the helix cd1, the ef linker (including the highly conserved sequence PEWY and the helix ef), and the amino-terminal end of helix F. Residues P271, F275, and M125 of cytochrome *b* and H161 of the Rieske protein, which has



**Figure 2** Inhibitor-binding sites. The electron-density maps (a, b) are from *bc*<sub>1</sub> crystals containing the inhibitors. a, Antimycin-binding site; electron-density map contoured at 0.7 Å. b, Stigmatellin-binding site; density contoured at 0.8 (blue) and 5.0 (orange) for the iron-sulphur cluster. The backbone of cytochrome *b* is in red and that of the iron-sulphur protein is in magenta. c, Schematic drawing of the secondary structure of cytochrome *b*. Given the number of a residue in the chicken sequence (used here), the number of the corresponding residue in the *Saccharomyces cerevisiae* sequence (conventionally used for alignment) is found by subtracting 2 if the number is less than 114. For residues 114 and later, the numbering is the same as in yeast. The number of the corresponding residue in the beef sequence is found by subtracting 1 from the number in the chicken sequence after the first five residues.

moved from its position in the native crystal (see below), are near the inhibitor. Residues 126–129 of helix C, and 140–147 of the linker cd, are also close by. In the native crystals, Y279 passes through the region at which we have modelled the stigmatellin head group, but in the stigmatellin-bound crystal Y279 has moved and interacts with R283 and with the Rieske backbone around C160. The electron density that we attribute to stigmatellin is strongly connected to the Rieske protein at the position of H161. This may represent a hydrogen bond responsible for holding the Rieske protein in its proximal position (see below) in the presence of myxothiazol. Formation of such a hydrogen bond between stigmatellin<sup>20</sup> or ubiquinol<sup>21,22</sup> at the Q<sub>o</sub> site and a Rieske cluster histidine has been previously suggested.

**The myxothiazol site.** Myxothiazol (not shown) binds in roughly the same place as stigmatellin, but is displaced slightly towards the centre of the membrane and the low-potential *b* haem group (*b<sub>L</sub>*). It is also close to P271, but whereas stigmatellin reaches outward from P271 toward the Rieske protein, myxothiazol and MOA-stilbene reach toward Y132 and F129 in helix C, in the vicinity of *b<sub>L</sub>*. This may be the site from which electron transfer from the ubisemiquinone to the cytochrome *b<sub>L</sub>* haem occurs.



**Figure 3** Structure of the intermembrane (external surface) domains of the chicken *bc<sub>1</sub>* complex. This is viewed from within the membrane, with the transmembrane helices truncated at roughly the membrane surface. Ball-and-stick models represent the haem group of cytochrome *c<sub>1</sub>*, the Rieske iron-sulphur cluster, and the disulphide cysteines of subunit 8. SU, subunit; cyt, cytochrome.

**Table 2 Distances between the Rieske iron-sulphur cluster and cytochromes *c<sub>1</sub>* and *b<sub>L</sub>***

| Crystal                                             | Distance (Å) from Fe <sub>2</sub> S <sub>2</sub> cluster to |                           | Designation (proximal/distal from haem <i>b<sub>L</sub></i> ) |
|-----------------------------------------------------|-------------------------------------------------------------|---------------------------|---------------------------------------------------------------|
|                                                     | Haem <i>b<sub>L</sub></i>                                   | Haem <i>c<sub>L</sub></i> |                                                               |
| Beef <i>P4</i> ,22 (from ref. 12)                   | 27.0                                                        | 31.0                      | Proximal                                                      |
| Chicken <i>P2</i> ,2,2 <sub>1</sub> (+stigmatellin) | 26.4                                                        | 31.6                      | Proximal                                                      |
| Chicken <i>P2</i> ,2,2 <sub>1</sub>                 | 34.3                                                        | 21.3                      | Distal                                                        |
| Beef <i>P6</i> ,22                                  | 34.9                                                        | 17.2                      | Distal                                                        |
| Beef <i>P2</i> ,1                                   | 35.1                                                        | 17.5                      | Distal                                                        |
| Rabbit <i>P6</i> ,22                                | 35.5                                                        | 19.1                      | Distal                                                        |

Iron peaks were located as peaks in electron density, calculated from averaged experimental phases and improved and extended by molecular averaging, except for chicken *P2*,2,2<sub>1</sub> in the absence of inhibitor, in which case Bivoet difference amplitudes were used with improved experimental phases retarded by 90°.

Arrangement of the intermembrane protein domains

Figure 3 shows a slab including the extrinsic domains in the intermembrane region of the chicken complex. The two cytochrome *c<sub>1</sub>* molecules contact each other through loops that surround an empty area around the two-fold axis. Subunit 8 (the 'hinge protein for formation of the cytochrome *c<sub>1</sub>*–*c* complex' of ref. 23) and the external ends of subunits 7 and 10 interact with cytochrome *c<sub>1</sub>* on the side away from the dimer interface. The hinge protein consists of a bent hairpin held by two internal disulphide bonds.

Structure of cytochrome *c<sub>1</sub>*

Cytochrome *c<sub>1</sub>* is one of the three redox-active proteins in the cytochrome *bc<sub>1</sub>* complex, but is incomplete in the beef complex structure<sup>12</sup>. The subunit is well ordered in our chicken crystals and the entire polypeptide can be traced. Its extrinsic domain forms a wedge-like structure containing the haem group, with a C-terminal transmembrane anchor next to helix E of cytochrome *b*. Figure 4 compares the backbone-folding patterns of cytochrome *c<sub>1</sub>* and mitochondrial cytochrome *c*, the prototype of Ambler's class I cytochromes *c*<sup>24</sup>. Cytochrome *c* has five helical segments, α1–α5.



**Figure 4** The structures of cytochrome *c<sub>1</sub>* and cytochrome *c*. Top, the ribbon diagram of mitochondrial cytochrome *c* with the open corner of the C pyrrole of the haem group facing the viewer, and the haem propionates directed downwards. Bottom, our current structure of cytochrome *c<sub>1</sub>*, rotated to put the common features between the two cytochromes in the same orientation. Corresponding segments of each cytochrome are drawn with the same colour. Helices labelled α1, α3, and α5 correspond to similarly labelled and coloured helices in cytochrome *c*, whereas those labelled α2\* and α6\* have no counterpart in cytochrome *c*.



Three helices ( $\alpha 1$ ,  $\alpha 3$  and  $\alpha 5$ ), which are conserved in class I cytochromes in general, are present in cytochrome  $c_1$  and occupy the same positions relative to each other and to the haem. Conserved aromatic residues involved in interactions between  $\alpha 1$  and  $\alpha 5$  (F10 and Y97 in mitochondrial cytochrome) are present as Y33

and F189, respectively. The tripeptide PNL, starting at residue 30, is conserved in mitochondrial cytochromes  $c$ . The proline carbonyl accepts a hydrogen bond from N $^{\delta}$  of the histidine haem ligand and the leucine provides a hydrophobic environment for the haem ring. This aligns with the tripeptide PDL that begins at residue 111 of cytochrome  $c_1$ . It is conserved in all cytochromes  $c_1$  except that of *Rhodobacter sphaeroides*, which, barring a sequencing error, has the tripeptide ADL. These similarities justify inclusion of cytochrome  $c_1$  as a class I cytochrome.

In mitochondrial cytochromes  $c$  the pyrrole C corner of the haem group is exposed at the 'front' face, where electron transfer is thought to take place. This corner is also exposed in our cytochrome  $c_1$  structure. The exposed C corner of the haem is surrounded by three regions of the protein, consisting of residues 36–41 (corresponding to cytochrome  $c$  13–18, 'fingerprint' region), the side chain of Y95 and residues 104–106 (helix 2'; no corresponding residues in cytochrome  $c$ ), and residues 158–163 (containing the haem ligand M160 and corresponding to cytochrome  $c$  residues 77–82).

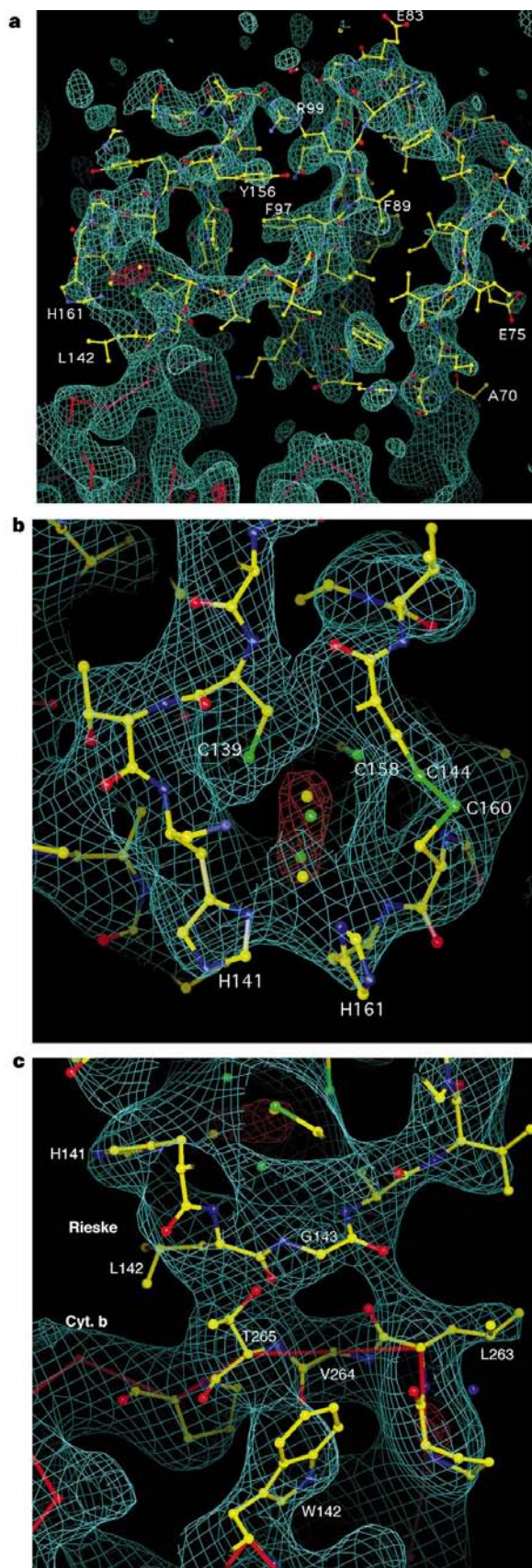
Major differences between cytochromes  $c$  and  $c_1$  are the result of insertions or deletions in loop regions. For example, bovine cytochrome  $c_1$  has an N-terminal extension of 24 residues before helix  $\alpha 1$ , whereas bovine cytochrome  $c$  has a one-residue extension. In cytochrome  $c_1$  this extension interacts with subunit 8, the hinge protein. After helix  $\alpha 1$  and the 'fingerprint' haem-binding stretch, CXXCH which are similar in the two cytochromes, cytochrome  $c_1$  has a long insertion (residues 42–109 of cytochrome  $c_1$  replace residues 18–28 of cytochrome  $c$ ). This expanded loop includes a region implicated in cytochrome  $c$  binding<sup>25</sup> and the dimer contact with cytochrome  $c_1$  in the other monomer seen in Fig. 3. Another insertion is found between the methionine haem ligand and helix 5; the six residues 81–86 in cytochrome  $c$  correspond to eighteen residues (161–178) in the  $c_1$  cytochromes. This region has also been implicated in cytochrome  $c$  binding<sup>26</sup>. The end of helix  $\alpha 5$  is the C terminus of cytochrome  $c$  but the transmembrane helix  $\alpha 6'$  is found after  $\alpha 5$  cytochrome  $c_1$ .

There is a second exposure of the haem on the A–D edge: the long loop present in cytochromes  $c$  and  $c_2$ , corresponding to residues 41–58 in tuna cytochrome  $c$ , is absent in cytochrome  $c_1$ . This results in exposure of the haem propionates to the surface. As described below, this edge is within electron-transfer distance of the iron–sulphur cluster in some crystals, indicating that this may be the pathway for reduction by the iron–sulphur protein.

### The Rieske iron–sulphur protein

Another of the three functionally important redox-active subunits of the cytochrome  $bc_1$  complex, the Rieske iron–sulphur protein, is missing in the structure of the tetragonal beef crystal<sup>12</sup>. Electron densities in the region of the globular extrinsic domain of this protein in our crystals are weaker than those in the rest of the structure, but are present and recognizable (Fig. 5). The backbone density is completely connected only when contoured at 1  $\sigma$  or lower, whereas the cytochrome  $b$  backbone in the transmembrane helices was continuous even when contoured at 3  $\sigma$ . However, the density was good enough to unambiguously locate the known structure of the soluble domain of the Rieske protein<sup>27</sup>.

◀ **Figure 5** The Rieske iron–sulphur protein. The electron-density maps are from improved experimental phases, contoured at 1  $\sigma$  (blue) and 5  $\sigma$  (red). The atomic model of the soluble domain of the iron–sulphur protein is from the coordinates of the protein database entry 1RIE (ref. 27), positioned as described in the text. **a**, A slab through the protein, including the iron–sulphur cluster (orange net) and the connection to the transmembrane helix (around residue 70). **b**, A close-up view of the electron density around the iron–sulphur cluster (orange net) of the Rieske protein. **c**, Contact of the Rieske protein (in the haem  $b$  distal conformation) with cytochrome (cyt)  $b$ . The model of cytochrome  $b$  (red C $\alpha$  backbone plus ball-and-stick models for residues W142 and L263 to P266) is from our coordinates.



As predicted from hydropathy plots and molecular-engineering results<sup>28,29</sup>, the iron–sulphur protein has a membrane-spanning helical segment near the N terminus. This was removed by proteolysis in preparing the soluble domain for structure determination<sup>27</sup>. In our electron-density map (Fig. 5a), the electron density continues before residue 70 where the model starts, and connects to a transmembrane helix. The transmembrane helix is well ordered.

The N-terminal 24 residues are on the matrix side, and interact with subunit 1. Residues 25–62 form a transmembrane helix, and are close to the transmembrane helices of subunit 10 and cytochrome *c*<sub>1</sub> (and, in the mammalian crystals, the putative subunit 11). The transmembrane helix is slightly curved and highly slanted. It passes through the membrane at an angle of about 32° to the two-fold axis, which is assumed to be perpendicular to the membrane. This high degree of tilt accounts for the length of the transmembrane helix (37 residues), which previously led to suggestions of two transmembrane helices for the Rieske protein<sup>29</sup>.

Residues 60–66 are in close contact with both cytochrome *b* subunits in the dimer, whereas residues 67–73 provide a flexible 'tether' connecting the extrinsic domain of the Rieske protein to its transmembrane helix. Figure 5b shows a close-up view of the iron–sulphur cluster region of the Rieske protein. Two histidine ligands, residues 141 and 161, are seen as bulges in the density at the tip of the protein. The iron atoms (orange net) are not individually resolved in this map.

### Swapping of extrinsic domain between monomers

Except for the transmembrane helix, only residues 141–143 of the iron–sulphur protein (one of the two loops that enclose the iron–sulphur cluster) contact with cytochrome *b* in the native chicken crystals. This contact (Fig. 5c) seems to involve interaction of residues L142 and G143 of the Rieske protein of one monomer with T265 and L263 of cytochrome *b* of the other monomer.

The extrinsic domain of the iron–sulphur protein has no contacts with the other extrinsic domains within a monomer in the native chicken crystals (Fig. 3). But the iron–sulphur cluster is close to the haem group of cytochrome *c*<sub>1</sub> of the other monomer within the complex dimer. As described below, this may provide the pathway for electron transfer between the iron–sulphur protein and cytochrome *c*<sub>1</sub>. Taking monomers to be as defined above, the iron–sulphur cluster of one monomer is in a position to transfer electrons with cytochromes *b* and *c*<sub>1</sub> of the other monomer.

The small number of contacts with the rest of the dimer probably accounts for the poor order of the Rieske extrinsic domain, and indicates that the domain may be mobile (also suggested in ref. 12). This mobility is restricted in one monomer of the chicken crystals,

and in the beef and rabbit hexagonal crystals, by interdimeric crystal contacts involving the extrinsic domain of the Rieske protein. The poor order of the extrinsic domain in the beef crystals and the large distance between the cluster and the haem of cytochrome *c*<sub>1</sub> indicated that this mobility may be required for function (ref. 12).

### Two conformations of the Rieske protein

Although the distances between the six haem-iron peaks of the dimer were the same (within experimental error) in all four crystal forms, the distance from the iron–sulphur cluster to an haem group varied by up to 5 Å in the different native crystals. From published distances between iron centres, it is clear that iron–sulphur clusters in the tetragonal beef crystals<sup>11,12</sup> and in any of our native crystals are positioned differently (Table 3). But when we treated the chicken cytochrome *bc*<sub>1</sub> complex with a saturating amount of stigmatellin before crystallization, the extrinsic domain of the iron–sulphur protein was found at a location different to that in native crystals, and the iron–sulphur cluster was in the same position as in the tetragonal beef<sup>12</sup> crystals. This movement can be simply and dramatically demonstrated using Bivoet difference maps constructed from diffraction data collected with X-ray wavelength near the iron absorption edge. Because of anomalous scattering by iron, the peaks in such maps indicate positions of irons in the complex: the three haem irons of the cytochromes and the iron–sulphur cluster of the Rieske protein. Bivoet difference maps are shown in Fig. 6. The peaks labelled Fe–S move closer to the haem groups of cytochrome *b* in the presence of stigmatellin. We call this the proximal conformation of the Rieske protein; the conformation in our native crystals is the distal conformation. The relative position of the iron–sulphur cluster in the chicken crystals containing stigmatellin is 16 Å from the position in the native chicken crystals, and 20 Å from that in the beef hexagonal crystals.

Stereo views of the two conformations of the Rieske protein, in the context of the entire *bc*<sub>1</sub> complex dimer and in isolation with cytochrome *b* and the haem of cytochrome *c*<sub>1</sub>, are shown in Fig. 1b, c. The two locations of the extrinsic domain of the Rieske protein are related by a rotation of 57° about an axis passing near residues 93 and 182 of the protein, perpendicular to the plane of the picture in Fig. 1c. The transmembrane helix and matrix-side portion are unchanged in the presence of stigmatellin. The coil consisting of residues 68–73 is stretched out in the presence of stigmatellin, allowing this end of the soluble domain to move farther from the membrane as the Fe<sub>2</sub>S<sub>2</sub> cluster on the other end moves closer. In a crystal containing bound stigmatellin and antimycin, the position of the iron–sulphur cluster was nearly the same as in the stigmatellin-bound crystal. In crystals containing only antimycin or

**Table 3 Structure-determination statistics: diffraction data for chicken *bc*<sub>1</sub> crystals**

|            | <i>d</i> <sub>min</sub> | Number of reflections | Unique reflections | Completeness (%)<br>( <i>I</i> > 1 $\sigma$ ) | <i>R</i> <sub>merge</sub> (%) | X-ray source* |
|------------|-------------------------|-----------------------|--------------------|-----------------------------------------------|-------------------------------|---------------|
| Native     |                         |                       |                    |                                               |                               |               |
| chn21      | 3.60                    | 279,119               | 70,363             | 76.3 (58.1)                                   | 18.6                          | RA            |
| chc01      | 3.10                    | 556,456               | 123,869            | 91.6 (80.5)                                   | 10.2                          | SSRL          |
| chm        | 3.01                    | 569,255               | 141,427            | 96.6 (74.2)                                   | 16.2                          | SSRL          |
| chb        | 2.95                    | 433,902               | 131,641            | 81.7 (51.2)                                   | 27.8                          | BNL           |
| Derivative |                         |                       |                    |                                               |                               |               |
| PIP†       | 3.50                    | 292,339               | 86,221             | 91.8 (76.3)                                   | 10.8                          | SSRL          |
| NSDMA‡     | 3.90                    | 425,028               | 67,109             | 99.7 (79.0)                                   | 12.4                          | RA            |
| TML02§     | 3.50                    | 203,105               | 61,380             | 65.2 (53.5)                                   | 17.1                          | RA            |
| TML03§     | 4.30                    | 110,201               | 38,103             | 74.4 (50.6)                                   | 21.6                          | RA            |
| HPDL¶      | 4.00                    | 129,187               | 48,715             | 78.4 (54.0)                                   | 20.2                          | RA            |
| Iridium    | 3.50                    | 177,303               | 68,522             | 71.7 (43.5)                                   | 13.4                          | RA            |
| TMLssrl§   | 3.50                    | 160,826               | 61,128             | 65.6 (43.9)                                   | 19.9                          | SSRL          |
| HPDLssrl¶  | 3.15                    | 350,204               | 93,367             | 71.6 (60.2)                                   | 19.3                          | SSRL          |

Each line corresponds to one data set collected from a single native or derivatized crystal of chicken cytochrome *bc*<sub>1</sub>.

\* X-ray sources and wavelengths are indicated by: RA, rotating anode (1.54 Å); SSRL, Stanford Synchrotron Radiation Laboratory BL7-1 (1.08 Å); BNL, Brookhaven National Laboratory X-12B (1.006 Å).

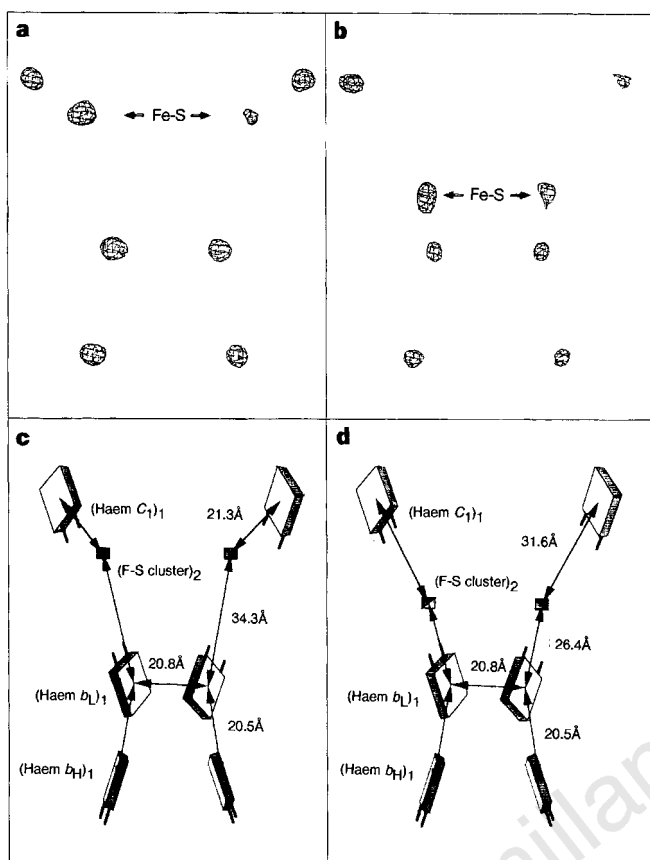
† Ethylenediamine platinum iodide<sub>2</sub>.

‡ *N*-(5-nitrosalicyl)-(S-decylmercuri)6-aminothiophenol, a putative antimycin analogue.

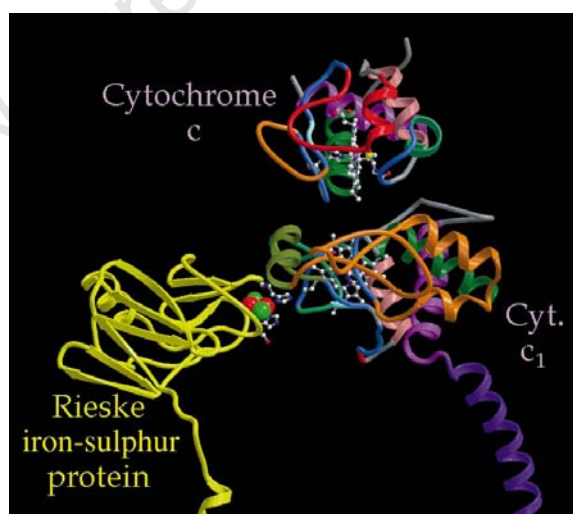
§ Trimethyl lead acetate (different concentrations and soaking times).

¶ Hexaphenyl di-lead.

|| Iridium carbonyl.



**Figure 6** Relative positions of the redox centres in the two different conformations of the  $bc_1$  complex dimer. **a, b**, Iron centres revealed by anomalous scattering near the iron edge. The net is a Bivoet difference map with X-ray wavelength 7.131 eV phased with experimental phases improved by averaging, and contoured at 4.5  $\sigma$ . **c, d**, Schematic drawing representing the cofactors. **a, c**, Results from a native crystal, with the iron-sulphur cluster of the Rieske protein in the distal position (from the low-potential haem group of cytochrome  $b$ ). **b, d**, Results from a crystal containing bound stigmatellin, with the cluster in the proximal position.



**Figure 7** Electron pathway through cytochrome  $c_1$  in a hypothetical complex of the  $bc_1$  complex with cytochrome  $c$ . The ribbon diagram shows the backbones of cytochrome  $c_1$ , cytochrome  $c$  (both with the same colour scheme as in Fig. 4) and the Rieske protein (yellow). The haem groups, the iron-sulphur cluster and surrounding residues are drawn as ball-and-stick models. The balls representing the iron-sulphur cluster (red and green) are enlarged for visibility. The position of the Rieske protein relative to cytochrome  $c_1$  is obtained from our beef hexagonal crystals.

myxothiazol, position of the iron-sulphur cluster was similar to that in the crystals without inhibitors.

In the proximal conformation, the iron-sulphur cluster ligand H161 of the Rieske protein is in H-bond distance of the occupant of the  $Q_0$  site (stigmatellin in our crystals (Fig. 2b), but, by inference, the electron donor ubiquinol *in vivo*), and in the distal conformation the iron-sulphur cluster is close to its electron acceptor, the haem group of cytochrome  $c_1$ . This suggests that the reaction mechanism for electron transfer in the cytochrome  $bc_1$  complex requires this dramatic conformational change, involving movement of the extrinsic domain of the iron-sulphur protein.

### Electron transfer to cytochrome $c_1$

In the native chicken crystals, the second loop of the Rieske protein enclosing the cluster (residues around H161) faces toward cytochrome  $c_1$ , approaching the haem propionates and residues 106 and 145 of cytochrome  $c_1$  (Fig. 3). As shown in Table 2, the Rieske protein is closer to cytochrome  $c_1$  in our two beef crystals. In these crystals there is electron-density contact at the 2  $\sigma$  level, between the Rieske protein around C160 (which forms a disulphide bond holding the cluster-binding loops together) and cytochrome  $c_1$  around G107 (between helix  $\alpha 2'$  and the haem-bracing P111). This electron density probably represents the configuration of the iron-sulphur protein during electron transfer to the cytochrome. Residue H161 of the Rieske protein provides one of the ligands to the  $Fe_2S_2$  cluster, and is 4.0 Å from an oxygen atom of haem propionate D and 8.2 Å from the edge of the haem  $\pi$ -bonded system at the C3D atom. From this distance (8.2 Å) we can calculate a rough rate of electron transfer from the iron-sulphur protein to cytochrome  $c_1$  of  $4.8-80 \times 10^6 \text{ s}^{-1}$ , assuming nonadiabatic electron tunnelling with reorganization energy of 0.7–1.0 electron volts and  $\Delta G^\circ$  near zero<sup>30,31</sup>. This is significantly faster than estimated rates for this reaction<sup>32</sup> so if the protein spends a small fraction of time in this conformation it could account for the reaction rate. In the native chicken crystals, this distance is 14.4 Å, which would give a rate of  $1.8-15 \times 10^3 \text{ s}^{-1}$  with the same assumptions. In the crystal containing stigmatellin, or the tetragonal beef crystals<sup>12</sup>, the shortest distance from the cluster or its ligands to the haem tetrapyrrole ring is 27 Å, giving (with the same assumptions) a rate of  $10^{-4} \text{ s}^{-1}$  and making it very unlikely that the enzyme could function in this single conformation.

### Model for electron transfer

Figure 7 shows a ribbon diagram of the extrinsic domains of the Rieske iron-sulphur protein and cytochrome  $c_1$ , as well as of cytochrome  $c$  bound to cytochrome  $c_1$  at a hypothetical site and orientation. The position of the iron-sulphur protein is that obtained from the beef  $P6_522$  crystals. This diagram illustrates the possibility of electron transfer into cytochrome  $c_1$  through the D propionate and out of cytochrome  $c_1$  through the C corner of the haem to cytochrome  $c$ . The distance between the two cytochromes is 10.2 Å, measured between atoms C2C of each haem group (the closest approach of the  $\pi$ -bonded systems). Assuming  $\Delta G^\circ$  near zero and reorganization energy  $\lambda$  in the range 0.7–1.0 gives electron-transfer rates in the range of  $0.6-5.1 \times 10^6 \text{ s}^{-1}$ . □

### Methods

**Purification and crystallization.** The cytochrome  $bc_1$  complex was purified from different vertebrate heart tissues nearly as described for the potato complex<sup>33</sup>. Mitochondria were prepared as described<sup>34</sup> and were solubilized using the detergent dodecyl maltoside. We isolated the complex from the extract by chromatography on DEAE Sepharose CL6B and further purified by size-exclusion chromatography on Sepharose CL6B<sup>13</sup>. The protein was concentrated to  $\sim 200 \mu\text{M}$  by ultrafiltration through an Amicon YM-100 membrane, precrystallized by mixing with 100 mM KMES pH 6.5 and 10% PEG-4000, and redissolved in 20 mM K-MOPS 7.5, 20  $\text{g l}^{-1}$  *n*-octyl- $\beta$ -D-glucopyranoside and 100 mM NaCl. Aliquots (5–20  $\mu\text{l}$ ) were mixed with an equal



volume of precipitant containing 20 mM KMES pH 6.7, 75 mM NaCl, 10% glycerol, and 6% PEG-4000, and were subjected to vapour diffusion against 30% glycerol.

To cocrystallize the  $bc_1$  complex with the high-affinity inhibitors antimycin, myxothiazol, and stigmatellin, we added inhibitor from an ethanolic solution (the final ethanol concentration was below 1% v/v) in a 1.5–2.0-fold molar ratio to the pooled fractions from the final column at a protein concentration of 5–10  $\mu$ M, before concentration and precrystallization as above.

**Cryogenic-data collection and reduction.** After crystallization was complete (5–30 days after setup), we added 20  $\mu$ l cryoprotectant containing 10 mM K-MES pH 6.7, 10 mM *n*-octyl- $\beta$ -D-glucopyranoside, 25% glycerol and 10% PEG-4000 to the solution containing the crystals from chicken complex, and changed the reservoir to 35% glycerol for further concentration of glycerol and PEG without increasing ionic strength. After this equilibration, or, in some cases, after further soaking in cryoprotectant consisting of 30% glycerol, K-MES and *n*-octyl- $\beta$ -D-glucopyranoside, crystals were frozen in liquid ethane or nitrogen, or in the cryogenic stream, and data were collected at 70–100 K. A suitable procedure for flash-freezing the beef and rabbit crystals has not yet been developed. We processed diffraction data by the programs DENZO and SCALEPACK<sup>35</sup>.

**Structure determination.** Data-collection statistics are summarized in Table 3. The chicken crystals were phased by isomorphous replacement and the resulting electron density was used to phase the other crystal forms by molecular replacement. Heavy-atom derivatives were first analysed using XtalView<sup>36</sup>. We used the RAVE package<sup>37</sup> for molecular averaging, map skewing, and rotation-translation-operator improvement. We used the CCP4 package<sup>38</sup> for final heavy-atom refinement and phase calculation (program MLPHARE) and for finding molecular replacement solutions (programs ALMN and TFCC). The phases were improved and extended to the resolution limit of the data by multicrystal and non-crystallographic symmetry averaging. During the phase improvement and extension process, correlation coefficients between the calculated electron-density map of the Rieske protein and our experimental electron density, monitored as a measure of the improvement of the maps, increased to 80–85% in the different data sets. The coefficient between subunits 1 and 2 increased to 40–48%.

Model building was done with the program O (ref. 39) and structures illustrated using this program or Molscript<sup>40</sup> and Raster3D (ref. 41).

All subunits of the  $bc_1$  complexes of vertebrates are expressed in the cytoplasm, except cytochrome *b*. Cytochrome *b*, which is expressed in mitochondria, has sequence identity of 74% between chicken and cows. As amino-acid sequences of the cytoplasmically expressed chicken subunits of the complex have not been reported, we used the sequences of the beef proteins for model building. Cytochrome *c*, a cytoplasmically expressed mitochondrial protein, has 89.5% identity between chicken and cows. Myosin light chain II of chicken is 91% identical to the human or mouse proteins.

**Location of iron centres from anomalous data.** Anomalous data at wavelength near the iron K absorption edge (7,131 eV) were collected for native and stigmatellin-containing crystals. Bivoet difference maps were made with coefficients of  $(F^+ - F^-)$  and improved experimental phases retarded by 90° to locate the iron centres.

**Electron density map calculations.** The electron-density maps were calculated using coefficients of  $(2F_o - F_c)e^{-i\Phi_c}$ , where the  $F_o$  values are from the experimentally determined intensities but the  $F_c$  and  $\Phi_c$  values are calculated from the previous map after multiple-crystal averaging. In the case of unobserved reflections,  $F_o$  was replaced by  $F_c$  as recommended<sup>42</sup>, resulting in coefficients of  $F_c e^{-i\Phi_c}$  for those terms. This 'fill-in' procedure was used both during averaging and, unless otherwise noted, in making the final maps used in the figures.

Received 12 December 1997; accepted 20 February 1998.

- Mitchell, P. Coupling of phosphorylation to electron and proton transfer by a chemiosmotic type of mechanism. *Nature* **191**, 144–148 (1961).
- Hinkle, P. C., Kumar, M. A., Resetar, A. & Harris, D. L. Mechanistic stoichiometry of mitochondrial oxidative phosphorylation. *Biochemistry* **30**, 3576–3582 (1991).
- Mitchell, P. Possible molecular mechanisms of the protonmotive function of cytochrome systems. *J. Theor. Biol.* **62**, 327–367 (1976).
- Crofts, A. R. in *The Enzymes of Biological Membranes*, Vol. 4 (ed. Martonosi, A. N.) 347–382 (Plenum, New York, 1985).
- Schägger, H., Link, T. A., Engel, W. D. & von Jagow, G. Isolation of the eleven protein subunits of the  $bc_1$  complex from beef heart. *Methods Enzymol.* **126**, 224–237 (1986).
- Schägger, H., Brandt, U., Gencic, S. & von Jagow, G. Ubiquinol-cytochrome-c reductase from human and bovine mitochondria. *Methods Enzymol.* **260**, 82–96 (1995).

- Weiss, H. & Leonard, K. Structure and function of mitochondrial ubiquinol:cytochrome c reductase and NADH:ubiquinone reductase. *Chemica Scripta* **27B**, 73–81 (1987).
- Yue, W. H., Zou, Y. P., Yu, L. & Yu, C. A. Crystallization of mitochondrial ubiquinol-cytochrome c reductase. *Biochemistry* **30**, 2303–2306 (1991).
- Kubota, T. et al. Crystallization and preliminary X-ray crystallographic studies of bovine heart mitochondrial cytochrome  $bc_1$  complex. *J. Mol. Biol.* **221**, 379–382 (1991).
- Berry, E. A., Huang, L.-S., Earnest, T. N. & Jap, B. K. X-ray diffraction by crystals of beef heart ubiquinol:cytochrome c oxidoreductase. *J. Mol. Biol.* **224**, 1161–1166 (1992).
- Yu, C.-A. et al. Crystallization and preliminary structure of beef heart mitochondrial cytochrome- $bc_1$  complex. *Biochim. Biophys. Acta* **1275**, 47–53 (1996).
- Xia, D. et al. Crystal structure of the cytochrome  $bc_1$  complex from bovine heart mitochondria. *Science* **277**, 60–66 (1997).
- Berry, E. A., Huang, L.-S., Shulmeister, V. M. & Kim, S.-H. A new form of crystal of bovine heart ubiquinol:cytochrome c oxidoreductase—determination of space group and unit cell parameters. *Acta Crystallogr. D* **51**, 235–239 (1995).
- Crofts, A. R. et al. in *The Phototrophic Prokaryotes—Proceedings of the IXth Int. Symp. on Phototrophic Prokaryotes*, Vienna, Sept. 1997 (eds Peschek, G. A., Loeffelhardt, W. & Schmetterer, G.) (Plenum, New York, London, Washington DC, Boston, in the press).
- Brandt, U., Yu, L., Yu, C. A. & Trumpower, B. L. The mitochondrial targeting presequence of the Rieske iron-sulfur protein is processed in a single step after insertion into the cytochrome  $bc_1$  complex in mammals and retained as a subunit in the complex. *J. Biol. Chem.* **268**, 8387–8390 (1993).
- Beaumont, K., Clough, J. M., Defraigne, P. J. & Godfrey, C. R. A. Fungicidal beta-methoxyacrylates—from natural products to novel synthetic agricultural fungicides. *Pestic. Sci.* **31**, 499–519 (1991).
- Clough, J. M. & Godfrey, C. R. A. Growing hopes. *Chemistry In Britain* **31**, 466–469 (1995).
- Sauter, H., Ammermann, E. & Roehl, F. in *Crop Protection Agents from Nature* (ed. Copping, L. G.) 50–81 (Royal Soc. Chem., Thomas Graham House, Cambridge, UK, 1996).
- Slater, E. C. The mechanism of action of the respiratory inhibitor, antimycin. *Biochim. Biophys. Acta* **301**, 129–154 (1973).
- Link, T. A. in *Frontiers of Cellular Bioenergetics: Molecular Biology, Biochemistry and Physiopathology* (eds Papa, S., Guerrieri, F. & Tager, J. M.) (Plenum, New York, London, in the press).
- Robertson, D. E., Daldal, F. & Dutton, P. L. Mutants of ubiquinol-cytochrome c2 oxidoreductase resistant to Qo site inhibitors: consequences for ubiquinone and ubiquinol affinity and catalysis. *Biochemistry* **29**, 11249–11260 (1990).
- Link, T. A. The role of the "Rieske" iron sulfur protein in the hydroquinone oxidation (Qp-) site of the cytochrome  $bc_1$  complex: the "proton-gated affinity change" mechanism. *FEBS Lett.* **412**, 257–264 (1997).
- Kim, C. H. & King, T. E. A mitochondrial protein essential for the formation of the cytochrome  $c_1$ -c complex. Isolation, purification, and properties. *J. Biol. Chem.* **258**, 13543–13551 (1983).
- Amler, R. P. Sequence variability in bacterial cytochromes c. *Biochim. Biophys. Acta* **1058**, 42–47 (1991).
- Stonehuerner, J. et al. Identification of the binding site on cytochrome  $c_1$  for cytochrome c. *J. Biol. Chem.* **260**, 5392–5398 (1985).
- Broger, C., Salardi, S. & Azzi, A. Interaction between isolated cytochrome  $c_1$  and cytochrome c. *Eur. J. Biochem.* **131**, 349–352 (1983).
- Iwata, S., Saynovits, M., Link, T. A. & Michel, H. Structure of a water soluble fragment of the 'Rieske' iron-sulfur protein of the bovine heart mitochondrial cytochrome  $bc_1$  complex determined by MAD phasing at 1.5 Å resolution. *Structure* **4**, 567–579 (1996).
- Van Doren, S. R., Yun, C.-H., Crofts, A. R. & Gennis, R. Assembly of the Rieske iron-sulfur subunit of the cytochrome  $bc_1$  complex in *Escherichia coli* and *Rhodospirillum rubrum* membranes independent of the cytochrome b and  $c_1$  subunits. *Biochemistry* **32**, 628–636 (1993).
- Link, T. A., Schägger, H. & von Jagow, G. in *Cytochrome Systems: Molecular Biology and Bioenergetics* (eds Papa, S., Chance, B. & Ernster, L.) 289–301 (Plenum, New York, 1987).
- Moser, C. C., Page, C. C., Farid, R. & Dutton, P. L. Biological electron transfer. *J. Bioenerg. Biomembr.* **27**, 263–274 (1995).
- Ding, H. et al. Ubiquinone pair in the Qo site central to the primary energy conserving reactions of cytochrome  $bc_1$  complex. *Biochemistry* **34**, 15979–15996 (1995).
- Crofts, A. R. & Wang, Z. How rapid are the internal reactions of the ubiquinol:cytochrome  $c_2$  oxidoreductase? *Photosynthesis Res.* **22**, 69–87 (1989).
- Berry, E. A., Huang, L.-S. & DeRose, V. Ubiquinol-cytochrome c oxidoreductase of higher plants. Isolation and characterization of the  $bc_1$  complex from potato tuber mitochondria. *J. Biol. Chem.* **266**, 9064–9077 (1991).
- Smith, A. L. Preparation, properties, and conditions for assay of mitochondria: slaughterhouse material, small scale. *Methods Enzymol.* **10**, 81–86 (1967).
- Otwinski, Z. in *Proceedings of the CCP4 Study Weekend: "Data Collection and Processing"* (eds Sawyer, L., Isaacs, N. & Bailey, S.) 56–62 (SERC Daresbury Laboratory, UK, 1993).
- McRee, D. E. A visual protein crystallographic software system for X11/XView. *J. Mol. Graph.* **10**, 44–46 (1992).
- Jones, T. A. in *Proc. CCP4 Study Weekend, Molecular Replacement* (eds Dodson, E. J., Glover, S. & Wolf, W.) 91–105 (SERC Daresbury Laboratory, UK, 1992).
- The SERC (UK) Collaborative Computing Project No. 4. The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallogr. D* **50**, 760–763 (1995).
- Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
- Kraulis, P. J. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
- Merritt, E. A. & Murphy, M. E. P. Raster3D version 2.0—a program for photorealistic molecular graphics. *Acta Crystallogr. D* **50**, 869–873 (1994).
- Rayment, I. Molecular replacement method at low resolution: optimum strategy and intrinsic limitations as determined by calculations on icosahedral virus models. *Acta Crystallogr. A* **39**, 102–116 (1983).

**Acknowledgements.** We thank T. Link and his co-workers for providing coordinates for the soluble domain of the Rieske iron-sulphur protein before the release from the Protein Data Bank; H. Bellamy for performing the XAFS scan and for advice on MAD data collection; S. Hong for preparing coordinate files for the inhibitors; T. Meyer for critically reading the manuscript; and L. Tong for advice on molecular averaging. This work was supported by the NIH (grants to E.A.B. and A.R.C.) and by the Office of Biosciences and Environmental Research, US Department of Energy (grant to S.-H.K.). The work was partially done at SSRL which is operated by the Department of Energy, Division of Chemical/Material Sciences. The SSRL Biotechnology Program is supported by the NIH Biomedical Resource Technology Program, Division of Research Resources.

Correspondence and requests for materials should be addressed to E.A.B. and S.-H.K. Atomic coordinates of the chicken  $bc_1$  complex have been deposited in the Brookhaven Protein Database for release in May 1998 (accession number 1BCC for the native chicken structure and 3BCC for the stigmatellin-antimycin-inhibited chicken structure).

# Outflow–infall interactions as a mechanism for terminating accretion in protostars

T. Velusamy & W. D. Langer

MS 169-506, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA

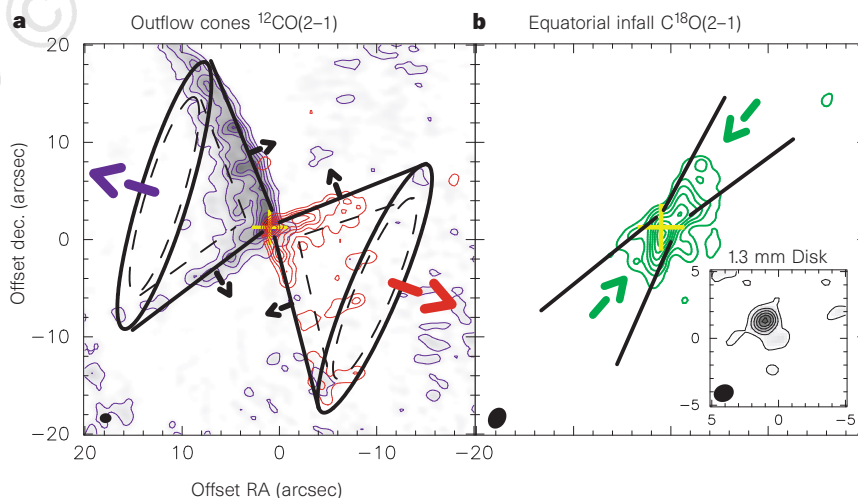
The formation of stars begins with the collapse of a dense interstellar cloud core to a protostar surrounded by a disk of gas and dust. Material in the envelope of the cloud core falls inwards to feed further growth of the protostar and its accretion disk. At some point during the accretion phase, an outflow of gas begins along the disk's rotation axis. Outflows have been studied in a large number of sources<sup>1</sup>, and recently it has become possible to study infall (and outflow) very close to the star<sup>2–8</sup>. But the possible interaction between these flows and its effect on the mass of the disk and the young star remain uncertain. Here we present observational evidence for an interaction between infalling and outflowing molecular gas. The opening angle of the outflow cone is largest near the star, indicating a widening of the outflow with time. Outside the lobes of the outflowing gas we see a narrow, disk-like region that is infalling. We suggest that the widening of the outflow may isolate the disk from further infall.

At the onset of the bipolar jet phase, the force of the resulting molecular outflow is sufficient to reverse the infall and sweep out material along the poles within the bipolar outflow cones. Therefore, in time the outflow could greatly modify the infall geometry. To identify the signatures of their time evolution we need high spatial- and velocity-resolution images of both infall and outflow structures. The high-velocity blue- and red-shifted outflow lobes are

detectable primarily in the <sup>12</sup>CO line wings. At the line centre velocities <sup>12</sup>CO emission is optically thick, and hence the denser inner regions must be probed in the optically thinner C<sup>18</sup>O isotope. Here we present interferometric images, obtained with the Owens Valley Radio Observatory Millimeter Array, of IRS1 in Barnard 5 (B5) which elucidate the structure of the outflow, infall and disk components on very small scales. IRS1 is a relatively young low-mass protostar embedded in a dense core in B5, located at a distance of 350 pc from Earth, with an outflow<sup>9</sup> and jet<sup>10</sup>. At an angular resolution of ~5", the CO outflow is biconical with a wide opening angle (ref. 2).

Our maps in Fig. 1, with angular resolution of 1", illustrate the first direct evidence (to our knowledge) of a very-wide-angle parabolic-shape molecular outflow close to the star, accompanied by infall in an equatorial disk. The high-velocity outflow originates in bipolar cones which can be traced deep down to within 0.4" from the star. The C<sup>18</sup>O(2–1) emission is confined to a narrow region, consistent with a gas disk perpendicular to the jet's outflow axis. The most remarkable feature revealed in these maps is the geometrical shapes of, and clear demarcation between, the <sup>12</sup>CO(2–1) and C<sup>18</sup>O(2–1) emissions. The disk, observed in the 1.3-mm continuum which best traces the inner pre-protoplanetary disk, is very compact with a size of ~400 AU. In contrast, the gaseous infall disk, which continues to fuel the growth of the protoplanetary disk (and protostar), is large: ~3,500 AU × 1,800 AU.

The outflow cones are seen best in the ±3 km s<sup>–1</sup> velocity emission (Fig. 2b). The integrated emission of the blue lobe (Fig. 3) shows in greater detail the vertex and lobe structure. Taken together our images show: (1) the CO outflow lobes show bipolar conical structure, and their vertices lie symmetrically on either side of the star within 150 AU. (2) The shape of the outflow lobe is parabolic near the vertices. (3) The bipolar outflow has evacuated a biconical cavity whose limb-brightened walls appear to be pushing and compressing the ambient and/or infalling gas. The cones have a sharp outer boundary and thin walls (≲500 AU). The hollow interior



**Figure 1** An overview of the outflow–infall interaction in IRS1 in B5. The observations were made in 1996–97 using the six-element array in both the high- and low-resolution configurations to obtain the best spatial frequency coverage and the highest resolution available, ~1" at 1.3 mm. **a**, Integrated <sup>12</sup>CO(2–1) emission in the high-velocity line wings traces the limb-brightened molecular outflow cones. The values of the first contours and contour intervals are 0.45 and 0.9 Jy km s<sup>–1</sup> per beam for the blue lobe, and 1.2 and 2.4 Jy km s<sup>–1</sup> per beam for the red lobe, respectively. The red and blue arrows denote red- and blue-shifted lobes, and the widening of the outflow cones is indicated by the short black arrows. **b**, The central low-velocity C<sup>18</sup>O(2–1) emission traces the infall region. The values of the first contour and contour interval are 0.24 Jy km s<sup>–1</sup> per beam. The

infall is confined to a flared disk ~30° wide, as indicated. The diameter of the disk is ~3,500 AU and its thickness varies from 850 AU at the centre to 1,700 AU at the edge. The inset in **b** shows the 1.3-mm continuum emission which traces the central pre-protoplanetary disk; the first contour and contour interval are 5 mJy per beam. In Figs 1–3 the offset position (0,0) corresponds to the IRAS position of IRS1 (ref. 23), and the crosses represent the continuum peak at right ascension (RA) 03 h 44 min 31.98(±0.009) s; declination 32° 42' 31.24(±0.06)" (equinox 1950), and the solid ellipses in the bottom left corner denote the size of the beam. The star's position inferred from the continuum disk is ~1.25" away from the IRAS position, and has an uncertainty <0.1" and, therefore, we adopt this for the star's position.

indicates a void in molecular gas. (4) The mean outflow velocity is nearly constant at  $\sim 6 \text{ km s}^{-1}$  across the lobes, so the structures seen within 1,000 AU are the result of the outflow generated in the recent past, within 1,000 yr. (5) Small-scale structures, about a few hundred AU, are seen in the outflow cones, which we suggest are evidence of episodic or variable outflow activity. (6) The momentum distribution along the outflow axis (estimated from the integrated intensity and the mean outflow velocity) peaks at the origin of the outflow, and decreases outwards. Although current models<sup>11–16</sup> can explain some of these observed outflow properties, none seem to explain all of them, especially the geometry both at the origin and at larger distances.

The parabolic shape of the outflow lobe structure near the vertex is best seen in the blue-shifted emission along the northeast direction (Fig. 3). We interpret this shape in terms of an increase in the outflow cone angle with time. In Fig. 3 we indicate the extent of the outflow in the cones at several epochs as inferred from the parabolic shape of the boundary, and assuming a mean outflow velocity of  $\sim 6 \text{ km s}^{-1}$  measured from  $^{12}\text{CO}$  line wings along the cone. Although the true ages of the protostar and outflow system as derived from the dynamical timescales may be uncertain<sup>17</sup>, the relative ages marked in Fig. 3, represent the chronological events. The geometrical shape at the vertex, which describes the outflow within the past 1,000 yr, is consistent with an opening angle of  $\sim 125^\circ$ , whereas at larger distances from the vertex it is narrower: for example, at 6,000 AU (originating about 6,000 yr ago) the corresponding opening angle is  $\sim 90^\circ$ . The difference implies a widening of the opening angle at a rate of  $\sim 0.006^\circ \text{ yr}^{-1}$ .

One way to discern the rotation and/or infall motions would be to look for velocity gradients along the major and minor axes (compare the results of Hayashi *et al.*<sup>18</sup>). The  $\text{C}^{18}\text{O}(2-1)$  velocity channel maps in Fig. 2a show a velocity gradient along the minor axis indicating that it is dominated by infall motions. The overall emission pattern in the  $\text{C}^{18}\text{O}(2-1)$  velocity maps is consistent with radial infall motions directed towards the star, within a flared disk,  $\sim 10'' \times 5''$ , which is inclined at  $\sim 45^\circ$  to the plane of the sky. The infall gas located out of the plane of the sky has a velocity shift  $\Delta V = V - V_{\text{LSR}}$  relative to the systemic local standard

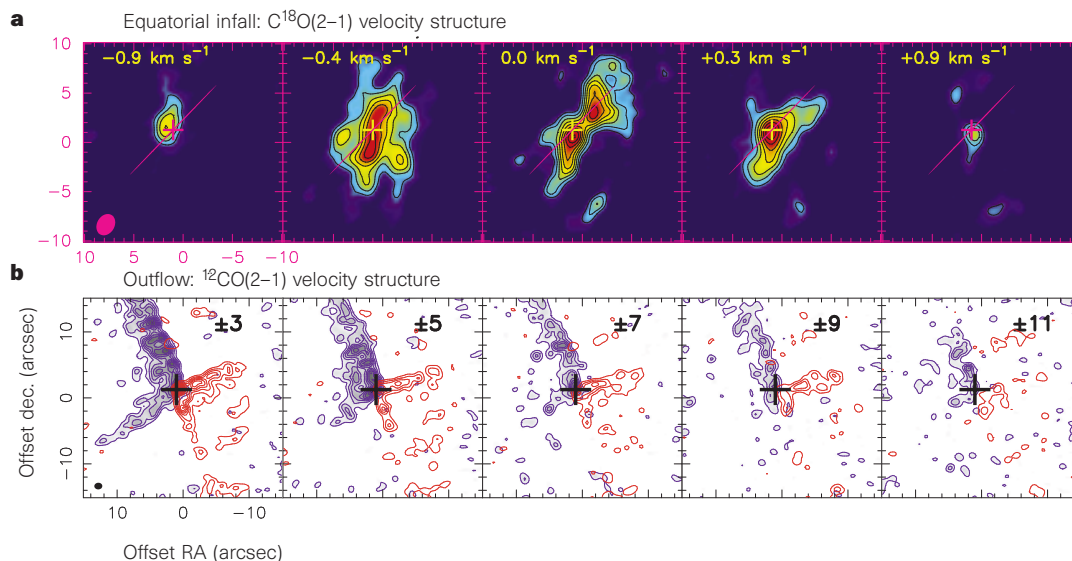
of rest (LSR) velocity  $V_{\text{LSR}} = 10.2 \text{ km s}^{-1}$ . The velocity shift  $\Delta V$  is  $< 0$  (towards the observer) on the far side, and  $\Delta V$  is  $> 0$  (away from the observer) on the near side, of the infall disk. The velocity gradient perpendicular to the equatorial plane (Fig. 2a) arises because of the inclination of the infall disk with respect to the plane of the sky. Because the infall velocity (approximately the speed of sound) increases towards the centre (that is, higher central density), the gas in the immediate vicinity of the protostar will show the largest infall velocities. Thus the compact emission centred near the star in the highest-velocity channel maps (outer panels in Fig. 2a) is consistent with infall. The channel map at the rest velocity (centre panel in Fig. 2a) traces the infalling gas which has a zero infall velocity component along the line of sight. In other words, this channel map shows the full spatial extent of the infall occurring in the plane of the sky. We note that the infall region is strictly confined to an equatorial zone which avoids the outflow (Fig. 1). The flared shape of the infall is reminiscent of the flared disk around HH30 (ref. 19).

At its present evolutionary stage, the protostar IRS1 in B5 has a flattened infall disk. It is reasonable to assume that this shape is the result of a continuous widening of the outflow lobes with time. It is also likely that outflow and infall are coupled to each other through the processes involved in generating the molecular outflow near the star. If IRS1 in B5 is young, as indicated by the dynamical age of outflow, we can estimate that, at the observed rate of outflow widening, it will take  $\sim 10^4$  yr to shut down the infall; therefore, the entire accretion phase lasts at least 20,000 yr after the onset of outflow. If it is much older, however, then it is harder to interpret the widening unless it simply arises due to a reduced pressure in the envelope. The timescales for accretion and the termination mechanism are important because they control the mass and subsequent evolution of a protoplanetary disk, the precursor to planet formation<sup>20</sup>.

Here we outline the time evolution of such an outflow–infall interaction qualitatively, assuming that IRS1 is young. In a jet-driven-through-shock model<sup>13,21</sup> for the outflow mechanism, the momentum transferred by the jet to the outflow per unit length along the outflow axis is related to the ambient density  $\rho_a$  by

**Figure 2** Selected channel maps of infall and outflow, with the velocities relative to  $V_{\text{LSR}}$  indicated in each panel.

**a**, The structure of  $\text{C}^{18}\text{O}(2-1)$  emission as a function of velocity. The first contour and contour interval are 0.8 and 0.2 Jy per beam, respectively. The southeast–northwest line through the centre represents the equatorial plane which is inclined to  $45^\circ$  to the plane of sky. We note that the velocity gradient perpendicular to the equatorial plane indicates radial infall motions within a flared disk. **b**,  $^{12}\text{CO}(2-1)$  velocity structure of red- and blue-shifted outflow lobes, at an angular resolution of  $1.2'' \times 1.0''$  shown at  $2 \text{ km s}^{-1}$  velocity spacing. The first contour and contour interval are 0.15 and 0.20 Jy per beam for the blue and red lobes, respectively.





$P_z \approx \pi r_j^2 v_j \rho_a \rho_j$ , where  $\rho_j$ ,  $r_j$  and  $v_j$  are respectively the density, radius and velocity of the jet. Therefore we expect the geometry and intensity of the outflow to depend on the location, extent and shape of the region of momentum transfer. Our data show that the largest momentum transfer is occurring closest to the star (less than a few hundred AU) as predicted by jet/shock-driven outflows<sup>21</sup> (and winds<sup>14</sup>). This result is consistent with the region of the highest density being most effective in transferring momentum. However, the density structure at this point is continuously modified owing to evacuation by outflow and accretion from infall. At the initiation of the jet the infall is spherical, the density distribution is broad and the outflow is narrow angled. As the outflow persists, material is swept out of the outflow cones, simultaneously reducing the accretion volume, by narrowing the extent over which infall occurs (Fig. 1). This interaction sets up a negative feedback in that the increase in the opening angle of the outflow cones creates a wider evacuated cavity above the disk–jet interface, and further narrows the equatorial infall. The resulting change in the density structure at the origin of outflow is likely to lead to a momentum transfer occurring in a wider angle. This process continues until the outflow opening angle approaches 180° and the infall is completely shut down.

The idea that the outflow opens up in time, thus blowing away the surrounding envelope and halting infall<sup>22</sup>, has remained a conjecture. The data presented here show an interesting possibility of outflow–infall interaction, leading to a change from spherical to an equatorial infall. We have shown that the opening angle of the outflow cones widens with time, which we interpret as a natural mechanism to stop the infall, and hence end the accreting phase in

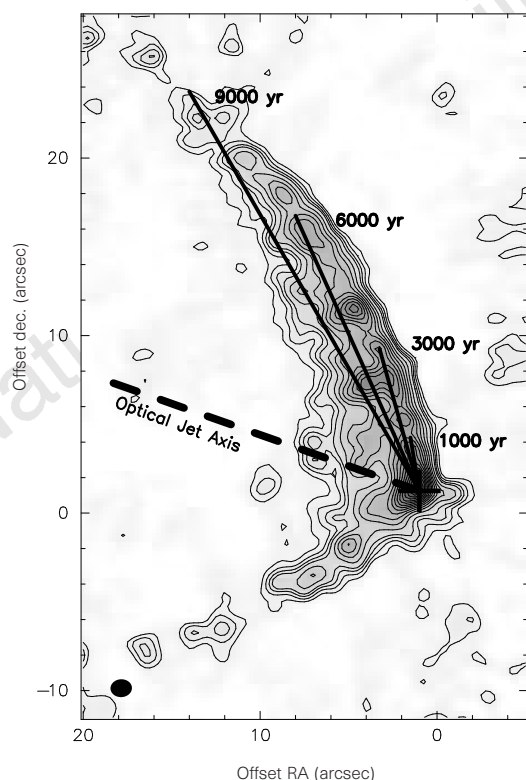
protostars. This observational evidence needs to be considered in theories of formation of stars and disks. □

Received 30 September 1997; accepted 14 January 1998.

1. Bontemps, S., Andre, P., Terebey, S. & Cabrit, S. Evolution of outflow activity around low-mass embedded young stellar objects. *Astron. Astrophys.* **311**, 858–872 (1996).
2. Langer, W. D., Velusamy, T. & Xie, T. IRS1 circumstellar disk and the origin of the jet and CO outflow in B5. *Astrophys. J.* **486**, L41–L44 (1996).
3. Chandler, C., Terebey, S., Barsony, M., Moore, T. J. T. & Gautier, T. N. Compact outflows associated with TMC-1 and TMC-1A. *Astrophys. J.* **471**, 308–320 (1996).
4. Chandler, C. & Sargent, A. I. The small scale structure and kinematics of B335. *Astrophys. J.* **486**, L29–L32 (1993).
5. Velusamy, T., Kuiper, T. B. H. & Langer, W. D. Imaging infall: CCS observations of protostellar envelope of B335. *Astrophys. J.* **451**, L75–L78 (1995).
6. Bachiller, R., Guilleaume, S. & Dutry, A. Jet driven molecular outflow in L1448-C. CO and continuum synthesis images. *Astron. Astrophys.* **299**, 857–868 (1995).
7. Gueth, F., Guilleaume, S., Dutry, A. & Bachiller, R. Structure and kinematics of a protostar: mm-interferometry of L1157. *Astron. Astrophys.* **323**, 943–952 (1997).
8. Ladd, C. J. & Fich, M. The structure and energetics of a highly collimated bipolar outflow: NGC2264G. *Astrophys. J.* **459**, 638–652 (1996).
9. Goldsmith, P. F., Langer, W. D. & Wilson, R. W. Molecular outflows, gas density distribution, and the effects of star formation in the dark cloud Barnard 5. *Astrophys. J.* **303**, L11–L15 (1986).
10. Bally, J., Devine, D. & Alten, V. A parsec-scale Herbig–Haro jet in Barnard 5. *Astrophys. J.* **473**, 921–928 (1996).
11. Stahler, S. The kinematics of molecular outflows. *Astrophys. J.* **422**, 616–620 (1994).
12. Raga, A. C., Cantó, J., Calvert, N., Rodríguez, L. F. & Torrelles, J. M. A unified stellar jet/molecular outflow model. *Astron. Astrophys.* **276**, 539–548 (1993).
13. Masson, C. R. & Chernin, L. M. Properties of jet-driven molecular outflows. *Astrophys. J.* **414**, 230–241 (1993).
14. Li, Z.-Y. & Shu, F. H. Magnetized singular isothermal toroids. *Astrophys. J.* **472**, 211–224 (1996).
15. Smith, M. D., Suttner, G. & Yorke, H. W. Numerical hydrodynamical simulations of jet-driven bipolar outflows. *Astron. Astrophys.* **323**, 223–230 (1997).
16. Nagar, N. M., Vogel, S. N., Stone, J. M. & Ostriker, E. C. Kinematics of the molecular sheath of the HH111 optical jet. *Astrophys. J.* **482**, L195–L198 (1997).
17. Parker, N. D., Padman, R. & Scott, P. F. Outflows in dark clouds: their role in protostellar evolution. *Mon. Not. R. Astron. Soc.* **252**, 442–461 (1991).
18. Hayashi, M., Ohashi, N. & Miyama, S. A dynamically accreting gas disk around HL Tauri. *Astrophys. J.* **418**, L71–L74 (1993).
19. Burrows, C. J. et al. Hubble Space Telescope observations of the disk and jet of HH30. *Astrophys. J.* **473**, 437–451 (1996).
20. Beckwith, S. & Sargent, A. I. Circumstellar disks and the search for neighbouring planetary systems. *Nature* **383**, 139–144 (1996).
21. Chernin, L. M. & Masson, C. R. Momentum distribution in molecular outflows. *Astrophys. J.* **455**, 182–189 (1995).
22. Shu, F., Najita, J., Galli, D., Ostriker, E. & Lizano, S. in *Protostars and Planets III* (eds Levy, E. H. & Lunine, J. I.) 3–45 (Univ. Arizona Press, Tucson, 1993).
23. Beichman, C. A. et al. Formation of solar type stars: IRAS observations of the dark cloud Barnard 5. *Astrophys. J.* **278**, L45–L48 (1984).

**Acknowledgements.** We thank Caltech's OVRO staff for their assistance and support. The Owens Valley Radio Observatory millimeter-wave array is supported by the NSF. This research was conducted at the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA.

Correspondence should be addressed to T.V. (e-mail: velu@rams.jpl.nasa.gov).



**Figure 3** Time evolution of the outflow cones is shown on the integrated intensity map of the blue-shifted outflow lobe. The direction of the jet as inferred from the optical data (which is not seen close to the star but only at parsec scales), is indicated. The first contour is 0.24 Jy km s<sup>-1</sup> per beam, and contour intervals are 0.24 for the first 6 contours and 0.48 for the rest. The extent of the outflow and direction of gas flow at the epochs shown are derived assuming a mean molecular outflow velocity of 6 km s<sup>-1</sup>. The widening of the opening angle with time is interpreted as a mechanism to stop infall.

## Discovery of a metastable $\pi$ -state in a superfluid $^3\text{He}$ weak link

S. Backhaus, S. Pereverzev, R. W. Simmonds, A. Loshak, J. C. Davis & R. E. Packard

Department of Physics, University of California, Berkeley, California 94720, USA

Under certain circumstances<sup>1,2</sup>, a superconducting Josephson junction can maintain a quantum phase difference of  $\pi$  between the two samples that are weakly connected to form the junction. Such systems are called ' $\pi$ -junctions' and have formed the basis of several experiments designed to investigate the much-debated symmetry of the order parameter of high-temperature superconductors<sup>3–8</sup>. More recently, the possibility that similar phenomena might occur in another macroscopic quantum system—a pair of weakly coupled Bose–Einstein condensates—has also been suggested<sup>9</sup>. Here we report the discovery of a metastable superfluid state, in which a quantum phase difference of  $\pi$  is maintained across a weak link separating two reservoirs of superfluid  $^3\text{He}$ . The existence of this state, which is the superfluid analogue of a superconducting  $\pi$ -junction, is likely to reflect the underlying ' $p$ -wave' symmetry of the order parameter of superfluid  $^3\text{He}$ , but a precise microscopic explanation is at present unknown.

Our experimental apparatus is essentially a container filled with superfluid B-phase  $^3\text{He}$  and partitioned into two reservoirs by a wall containing both a flexible membrane and a quantum weak link<sup>10</sup>, as shown schematically in Fig. 1. The weak link consists of a square array of 4,225 apertures, each of diameter 100 nm, spaced 3  $\mu\text{m}$  apart, and etched through a 50-nm-thick SiN membrane. The membrane, which is coated with a superconducting film, faces a displacement transducer<sup>11</sup> (based on a superconducting quantum interference device, SQUID) that can detect motion of the membrane as small as  $10^{-14}$  mHz<sup>-1/2</sup>. A nearby electrode can be used to move the membrane by the application of voltages between the electrode and the conducting film on its surface. The liquid is in contact with a nuclear demagnetization refrigerator that can cool the  $^3\text{He}$  to below 200  $\mu\text{K}$ .

The state of each reservoir can be described by a macroscopic quantum wavefunction, which has a well defined quantum phase  $\phi$  (ref. 12). The weak link can be characterized by a function  $I(\Delta\phi)$  relating the mass current  $I$  to the quantum phase difference  $\Delta\phi$  between the two reservoirs. The nature of  $I(\Delta\phi)$  at temperature  $T$  is determined by the ratio of the individual aperture size to the superfluid healing length  $\xi = \xi_0(1 - T/T_c)^{-1/2}$ , where  $\xi_0 = 65$  nm, and  $T_c = 0.929$  mK is the superfluid transition temperature.

Measurements of  $I(\Delta\phi)$  are carried out by analysing the time dependence of the position,  $x(t)$ , of the flexible membrane in response to various applied forces. From this record both  $I(t)$  and  $\Delta\phi(t)$  can be simultaneously and continuously obtained as follows. The instantaneous mass current through the array is given by:

$$I(t) = \rho A \frac{dx}{dt} \quad (1)$$

where  $A$  is the area of the membrane and  $\rho$  is the liquid density. The time dependence of  $\Delta\phi$  is given by the Josephson–Anderson phase evolution equation:

$$\Delta\phi(t) = \frac{-2m_3}{\rho\hbar} \int_0^t P(t)dt = \frac{-2m_3k}{\rho\hbar} \int_0^t x(t)dt \quad (2)$$

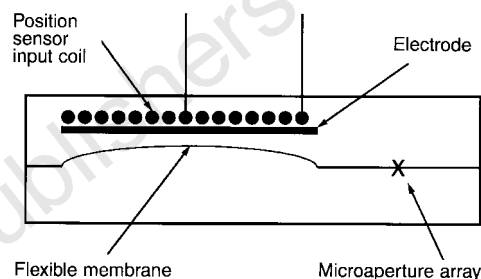
where  $k$  is the stiffness of the soft membrane (in Pa m<sup>-1</sup>),  $\hbar$  is Planck's constant,  $m_3$  is the atomic mass of  $^3\text{He}$ , and  $P$  is the pressure difference between the two reservoirs. We have previously<sup>10</sup> established an absolute pressure calibration for the constant  $k$  by correlating the measured Josephson frequency ( $\omega_1 = 2m_3P/\rho\hbar$ ), against the membrane displacement  $x$ . It is the result of that measurement that allows the absolute knowledge of the quantum phase difference  $\Delta\phi(t)$ .

The membrane has a normal mode of oscillation about  $\Delta\phi = 0$ , whose equations of motion are similar to that of a physical pendulum<sup>13</sup>. In prior measurements of this pendulum mode<sup>14</sup> we found that the full range of  $\Delta\phi$ , that is,  $\pi \rightarrow -\pi$ , was inaccessible when the temperature was below approximately  $0.48 T/T_c$  (where  $\xi$  is smaller than the aperture diameter). To probe greater values of  $\Delta\phi$ , we have now developed an experimental method that led to the discovery of the metastable superfluid state reported here. In our present experiment, we excite the oscillator by a few sine-wave cycles of an electrostatically applied, oscillating pressure. The excitation frequency matches the low-amplitude resonant frequency of the pendulum mode. The objective is to pump the pendulum mode to increasingly higher amplitudes to force the system to sample regions of  $\Delta\phi$  near  $\pm\pi$ .

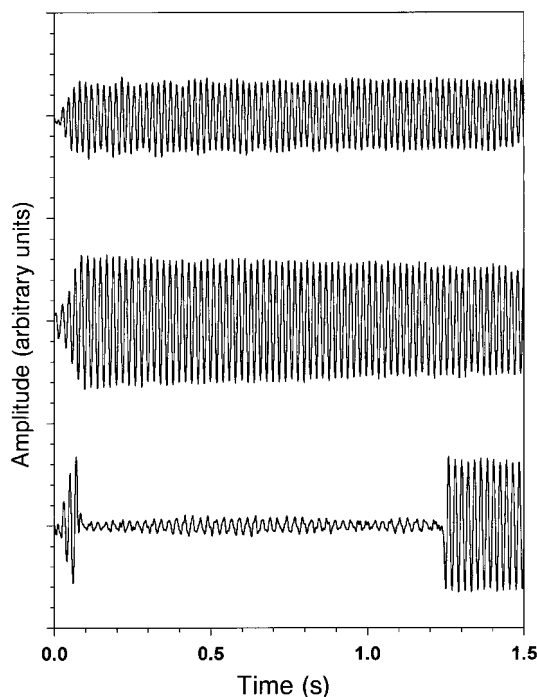
As shown in Fig. 2, a small excitation amplitude drives the oscillator to a small displacement amplitude. After the excitation ends, the oscillation decays slowly (top two traces) due to dissipation associated with the motion of the membrane. As either the number of cycles or the excitation level is increased, the oscillator achieves even larger amplitudes. Finally, an amplitude is reached where the oscillation collapses dramatically to a different mode with a lower amplitude and frequency (lowest trace). After many cycles in this new state, and without any further external drive, the oscillator

always makes a second transition back to a state with the same frequency and nearly the same amplitude as the initial state. Thus, during the time interval that the system remains in the low-amplitude mode, the kinetic energy in the original mode must be stored in some other degree of freedom.

The nature of this new state begins to be revealed by using equation (2) to convert  $x(t)$  into a plot of  $d(\Delta\phi)/dt$  versus  $\Delta\phi$ . As shown in Fig. 3a, during the excitation section of the experiment, the phase-difference oscillation is centred about  $\Delta\phi = 0$ . When the oscillation collapses, the measured quantum phase difference is found to make a transition to an oscillation about  $\Delta\phi = \pi$ . The smaller amplitude of the  $\Delta\phi$  oscillation around  $\pi$  is reflected in the smaller radius of the orbits about the new equilibrium. With dissipation, the amplitude of this oscillation can fall to below measurable levels and the system has been observed to remain in this state for more than 15 minutes.



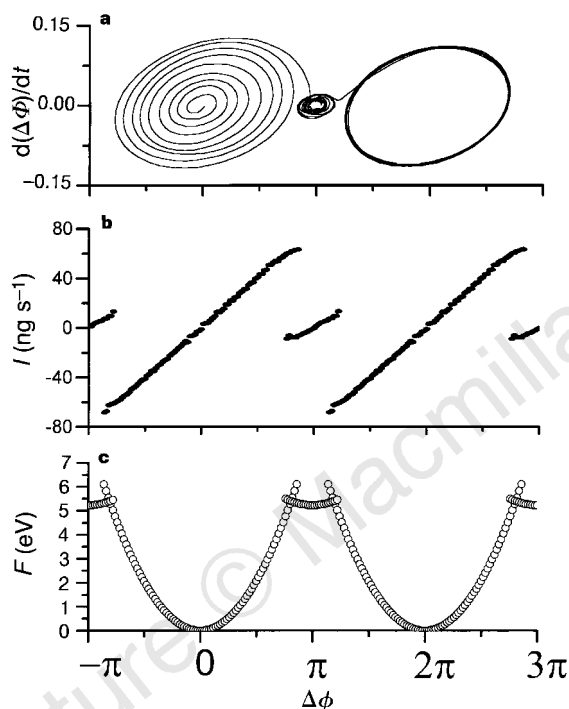
**Figure 1** A schematic diagram of the cell. A container filled with superfluid  $^3\text{He-B}$  is partitioned by a wall containing a flexible membrane and a weak link array.



**Figure 2** Oscillation of the membrane resulting from the application of 4 cycles of drive at the pendulum-mode resonant frequency, at  $0.37/T_c$ . The amplitude of the drive increases from top to bottom. The top two curves show a response similar to that of a high- $Q$  pendulum-like oscillator. The amplitude of motion increases while the drive is applied. After the drive is turned off, the oscillation decays slowly, due to the dissipation associated with the membrane motion. The bottom trace shows the departure from the behaviour of a simple oscillator. The drive level has been adjusted so that the oscillator makes the transition to a new state near the time when the drive is turned off. Without further excitation, the oscillator 'rings' freely in its new state until a random external perturbation causes the transition back to its original state. When the oscillator returns to its original state it regains the energy it had before the collapse.

The location of this new equilibrium has led us to name this state the ' $\pi$ -state'. By analogy, the equilibrium centred on  $\Delta\phi = 0$  is referred to as the 'zero-state'. When the oscillator makes the transition back to its original amplitude and frequency,  $\Delta\phi$  is found to switch back to oscillations about either  $\Delta\phi = 0$  or  $2\pi$ . Oscillations about  $2\pi$  are equivalent to the zero-state owing to the  $2\pi$  periodicity of  $I(\Delta\phi)$ . As the two modes of oscillation are about equilibrium points separated by  $\pi$ , we call these sudden changes in the mean phase difference ' $\pi$  phase slips'.

Using equations (1) and (2), we can compute  $I(\Delta\phi)$  over the full range of  $\Delta\phi$  directly from data obtained using our experimental technique. The result of such an analysis, shown in Fig. 3b, reveals that a new branch appears in the apparent current–phase relation. The previously unobservable portion of  $I(\Delta\phi)$  near  $\pm\pi$  is found to lie on the new branch with positive slope centred at  $\Delta\phi = \pi$ . The horizontal axis in Fig. 3 is the measured phase difference, and



**Figure 3** Oscillator characteristics. **a**, Trajectory of the oscillator in the  $d\Delta\phi/dt$  versus  $\Delta\phi$  plane at  $T/T_c = 0.28$ . During the drive section of the experiment, the oscillator moves in an increasing radius orbit centred on  $\Delta\phi = 0$ ; this increase is due to the energy input from the drive voltage. At the time the drive is turned off, the system makes a  $0 \rightarrow \pi$  transition and orbits about  $\Delta\phi = \pi$ . The oscillator rings in this state for many cycles (orbits) under the effects of light dissipation. At some time later, an external vibration provides enough energy to put the system into an orbit centred on  $2\pi$ . This state is equivalent to the state at  $\Delta\phi = 0$ . **b**, The measured current–phase relation for the microaperture array at  $0.28T/T_c$ . The vertical axis is the total mass current through the array. The  $\pi$ -branch appears only at temperatures below  $0.48 T/T_c$ . At higher temperatures  $I(\Delta\phi)$  approaches the sine-function behaviour of a classic Josephson weak link<sup>14</sup>. The maximum value of  $I$  at a given value of temperature varies between cool-downs through  $T_c$ . This is presumably because of trapped circulation in the array. **c**, The total energy stored in the weak link,  $F$ , as a function of  $\Delta\phi$  at  $0.28T/T_c$ . The energy is calculated from the data in Fig. 2b using equation (3). The metastable energy minimum  $F(\pi)$  is raised above  $F(0)$  by the measured energy jump at the transition  $0 \rightarrow \pi$ . We note that the horizontal axis represents the phase difference across the complete array and includes a hydrodynamic contribution due to the fluid entering and leaving the array. If this contribution is removed there is no longer an apparent degeneracy in either the energy–phase diagram, or the current–phase diagram.

includes a contribution from the hydrodynamic inductance near the apertures. It is this component which leads to the multivalued appearance of Fig. 3b and c. If that component is removed, the true  $I(\Delta\phi)$  and  $F(\Delta\phi)$  are single-valued. However, the dynamics of the system, which are the subject of this work, are determined by the overall phase difference as shown in the figure.

The implications of this more complete knowledge of  $I(\Delta\phi)$  are apparent when one considers the potential energy stored in the weak link. The free energy associated with the weak link is given by<sup>15</sup>:

$$F(\Delta\phi) = \frac{\hbar}{2m_3} \int_0^{\Delta\phi} I(\Delta\phi') d(\Delta\phi') \quad (3)$$

The results of calculating  $F$  from the experimental data are shown in Fig. 3c. The vertical position of  $F(\pi)$  is determined from the apparent energy decrement measured when the zero-state collapses to the  $\pi$ -state.

On referring to Fig. 3c, the interpretation of the motion of the oscillator is clear. Before the excitation is turned on, the system is stationary in its lowest-energy state at  $\Delta\phi = 0$ . When the oscillatory drive is switched on, it transfers energy to the system causing it to oscillate in the potential-energy minimum centred at  $\Delta\phi = 0$ . If, at the time when the drive is switched off, the amplitude of oscillation has reached the critical value of  $\Delta\phi$  where the two states intersect, the transition  $0 \rightarrow \pi$  may occur. With the drive off, and in the presence of only light dissipation, the system oscillates in this new, higher, local minimum centred at  $\pi$ . Although the kinetic energy of the fluid motion in the  $\pi$ -state may eventually dissipate, the energy difference between the two minima is stored in some other degree of freedom of the system and can be released when a transition occurs back to the zero-state. The transition  $\pi \rightarrow 0$  occurs when a perturbation, due to vibrations of the cryostat, contributes enough energy so that its oscillation amplitude becomes large enough to allow the transition.

The time the system spends in this new state is not deterministic, although some general observations can be made. At the lowest temperatures, the system can spend many minutes in this state. While at temperatures near  $0.48 T/T_c$  where the state first appears, the time is usually limited to less than 100 ms. The metastability is an unusual feature of our system which will require further research.

A number of situations have been proposed whereby a superconducting Josephson junction can be a ' $\pi$ -junction'. An early suggestion was for a  $\pi$ -junction in which magnetic scattering of electrons in the junction region flips the spins<sup>1</sup>. The idea of a  $\pi$ -junction was later connected to the symmetry of unconventional superconductors with the suggestion that a closed superconducting loop containing a sample of  $p$ -wave superconductor, which is connected to the ordinary superconductor by two Josephson junctions, could contain a  $\pi$ -junction<sup>2</sup>. Elaborations of these ideas have now achieved wide usage in the design of experiments to study the order-parameter symmetry of high-temperature superconductors. These experiments include: (1) SQUIDs formed with junctions on different  $a$ – $b$  edges of a crystal of high-temperature superconduction<sup>3,4,15,16</sup>, (2) studies of tricrystal rings which have junctions formed by grain boundaries in specific relative orientations<sup>6–8</sup>, and (3) junctions between a conventional superconductor in tunnelling contact with both sides of a twin boundary in a high-temperature superconductor<sup>17</sup>. Experiments using these geometries have made significant contributions to knowledge of the symmetry of high-temperature superconductors, and illustrate the importance of the idea of the  $\pi$ -junction.

Recently theoretical studies of Josephson effects between two weakly connected Bose–Einstein condensate (BEC) samples have also revealed the potential for the existence of  $\pi$ -states, which are created when the populations of the two BECs are strongly perturbed by the Josephson current<sup>9</sup>.

Superfluid  $^3\text{He}$  is characterized by a  $p$ -wave order parameter<sup>18</sup>. Although there are no crystal axes in this system, the symmetry of



the order parameter may allow locally stable phase differences of  $\pi$  to exist across a weak link between two containers of superfluid  $^3\text{He}$ . The depth of the observed energy well at  $\Delta\phi = \pi$  indicates the existence of a strong term in the  $I(\Delta\phi)$  function which is proportional to  $\sin(2\Delta\phi)$ . The microscopic reason for the existence of this term is so far unknown, and potential explanations in terms of local textures near the apertures, trapped circulation, and internal degrees of freedom of the superfluid will provide fruitful ground for further research.

**Note added in proof:** It has recently been pointed out<sup>19</sup> that there is a theoretical prediction of a second branch of  $I(\Delta\phi)$  at  $\Delta\phi = \pi$  in superfluid  $^3\text{He}$ , when the  $n$ -vector fields are antiparallel on opposite sides of the weak link. □

Received 20 November 1997; accepted 16 February 1998.

1. Bulaevskii, L. N., Kuzii, V. V. & Sobyannin, A. A. Superconducting system with weak coupling to the current in the ground state. *JETP Lett.* **25**, 290–293 (1977).
2. Geshkenbein, V. B., Larkin, A. I. & Barone, A. Vortices with half magnetic flux quanta in “heavy fermion” superconductors. *Phys. Rev. B* **36**, 235–238 (1987).
3. Wollman, D. A., Van Harlingen, D. J., Lee, W. C., Ginsberg, D. M. & Legett, A. J. Experimental determination of the superconducting pairing state in YBCO from the phase coherence of YBCO-Pb dc SQUIDS. *Phys. Rev. Lett.* **74**, 797–801 (1993).
4. Brawner, D. A. & Ott, H. R. Evidence for an unconventional superconducting order parameter in  $\text{YBa}_2\text{Cu}_3\text{O}_{6.9}$ . *Phys. Rev. B* **50**, 6530–6533 (1994).
5. Mathai, A., Gim, Y., Black, R. C., Amar, A. & Wellstood, F. C. Experimental proof of a time-reversal-invariant order parameter with a  $\pi$  phase shift in  $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ . *Phys. Rev. Lett.* **74**, 4523–4527 (1995).
6. Tsuei, C. C. *et al.* Pairing symmetry and flux quantization in a tricrystal superconducting ring of  $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ . *Phys. Rev. Lett.* **73**, 593–597 (1994).
7. Tsuei, C. C. *et al.* Symmetry of the order parameter of the high- $T_c$  superconductor  $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$ . *Nature* **373**, 225–229 (1995).
8. Tsuei, C. C. *et al.* Pairing symmetry in single layer tetragonal  $\text{Ti}_2\text{Ba}_2\text{CuO}_{6+\delta}$  superconductors. *Science* **271**, 329–332 (1996).
9. Smerzi, A., Fantoni, S., Giovanazzi, S. & Shenoy, S. R. Quantum coherent atomic tunneling between two trapped Bose-Einstein condensates. *Phys. Rev. Lett.* **79**, 4950–4953 (1997).
10. Pereverzev, S. V., Backhaus, S., Loshak, A., Davis, J. C. & Packard, R. E. Quantum oscillations between two weakly coupled reservoirs of superfluid  $^3\text{He}$ . *Nature* **388**, 449–451 (1997).
11. Paik, H. J. Superconducting tunable-diaphragm transducer for sensitive acceleration measurements. *J. Appl. Phys.* **47**, 1168–1178 (1976).
12. Tilley, D. R. & Tilley, J. *Superfluidity and Superconductivity* (Hilger, New York, 1990).
13. Monien, H. & Tewordt, L. Theory of Josephson flow oscillations in superfluid  $^3\text{He}$ -B. *J. Low Temp. Phys.* **62**, 277–300 (1986).
14. Backhaus, S., Pereverzev, S. V., Loshak, A., Davis, J. C. & Packard, R. E. Direct measurement of the current–phase relation of a superfluid  $\text{He}$ -3-B weak link. *Science* **278**, 1435–1438 (1997).
15. Barone, A. & Paterno, G. *Physics and Applications of the Josephson Effect* Section 1.6 (Wiley, New York, 1982).
16. Van Harlingen, D. J. Phase-sensitive tests of the symmetry of the pairing states in the high-temperature superconductors—Evidence for  $d_{x^2-y^2}$  symmetry. *Rev. Mod. Phys.* **67**, 515–535 (1997).
17. Kouznetsov, K. A. *et al.*  $c$ -axis Josephson tunneling between  $\text{YBa}_2\text{Cu}_3\text{O}_{7-\delta}$  and Pb: direct evidence for mixed order parameter symmetry in high- $T_c$  superconductors. *Phys. Rev. Lett.* **79**, 3050–3053 (1997).
18. Vollhardt, D. & Wolfe, P. *The Superfluid Phases of Helium-3* Section 7–2 (Taylor & Francis, New York, 1990).
19. Thuneberg, E. V. Two-dimensional Ginzburg–Landau simulation of the Josephson effect in superfluid  $^3\text{He}$ . *Europhys. Lett.* **7**, 441–446 (1998).

**Acknowledgements.** We thank D. Rokhsar, D.-H. Lee, S. Vitale, D. Weiss, S. Bahcall, R. Chiao, A. Smerzi and D. J. Van Harlingen for conversations. This work was supported in part by the National Science Foundation, the Office of Naval Research, the Packard Foundation, and the A.P. Sloan Foundation.

Correspondence and requests for materials should be addressed to J.C.D. (e-mail: jcdavis@physics.berkeley.edu).

## Phase-mapping of periodically domain-inverted $\text{LiNbO}_3$ with coherent X-rays

Z. H. Hu\*, P. A. Thomas\*, A. Snigirev†, I. Snigireva†, A. Souvorov†, P. G. R. Smith‡, G. W. Ross‡ & S. Teat§

\* Department of Physics, University of Warwick, Coventry CV4 7AL, UK

† ESFR, BP 220, F-38043 Grenoble, France

‡ Optoelectronics Research Centre, University of Southampton, Southampton SO17 1BJ, UK

§ Daresbury Laboratory, Warrington WA4 4AD, UK

A varying refractive index across a wavefront leads to a change in the direction of propagation of the wave<sup>1,2</sup>. This provides the basis for phase-contrast imaging of transparent or weakly absorbing materials with highly coherent X-ray beams<sup>3,4</sup>. Lattice distortions can also change the direction of propagation of a wave field

diffracted from a crystal. Here we report the use of this principle to effect phase-contrast imaging of the domain structure of a ferroelectric material, lithium niobate. A periodically domain-inverted structure for quasi-phase-matching of second-harmonic generation is created in this material, in which the direction of spontaneous polarization is sequentially inverted. Because of complex interactions during domain-inversion processing, this is accompanied by lattice distortions across the domain walls. These distortions split the diffracted wavefront of a beam of coherent X-rays from an advanced synchrotron source, giving rise to a pattern of interference that reflects the underlying pattern of lattice distortions. These results show that this phase-contrast imaging technique with sub-micrometre spatial resolution permits the non-destructive, highly sensitive phase-mapping of various structural defects and distortions introduced into materials during processing.

Interference and phase-contrast imaging at optical wavelengths can be conveniently realized by using optical lenses<sup>1</sup>. But this is not the case for hard X-ray wavelengths because the refractive index for X-rays differs from unity by only  $10^{-5}$  to  $10^{-6}$ . Therefore, methods which are free of lenses have been used for X-ray phase-contrast imaging such as the Bonse–Hart interferometer<sup>2</sup> that splits and subsequently recombines coherently-related X-ray beams to generate interference patterns. However, the stringent experimental conditions required with a conventional X-ray source have limited the applications of the phase-contrast imaging technique in research in condensed-matter physics, materials and biological sciences. This technique has recently aroused much interest because the high degree of coherence of intense X-ray beams provided by the European Synchrotron Radiation Facility (Grenoble) has made phase-contrast imaging possible in a convenient manner<sup>4,6,7</sup>.

We have explored the possibility of phase-contrast imaging of periodic domain inversion in ferroelectric nonlinear optical materials. Periodic domain inversion refers to the deliberate introduction of a periodic array of domains of alternating structural polarity into a polar crystal, displacing the cations and anions of the structure by applying an external electric field such that the resulting spontaneous polarization is reversed with respect to the original. This generates a spatially periodic modulation of the nonlinear coefficient along the direction of wave propagation that enables quasi-phase-matching of any frequency-doubling interaction within the transparency range of the crystal according to the choice of the period of modulation<sup>8</sup>. Lattice distortions occur during domain-inversion processing because of complex interactions between the physical effects taking place, for example, the converse piezoelectric and/or pyroelectric effects, and the dynamic process of domain inversion<sup>9–11</sup>.

The crystal used for our studies is  $\text{LiNbO}_3$ . A domain-inverted structure of period  $30\text{ }\mu\text{m}$  was fabricated in a  $\text{LiNbO}_3$  sample with a thickness of  $200\text{ }\mu\text{m}$  via electric-field poling. A synchrotron-radiation X-ray experiment was performed at the ESRF optics beamline (BL10/D5) which is on a bending magnet. The experimental configuration is schematically shown in Fig. 1.

Sets of synchrotron X-ray images of the 006 symmetric reflection are shown in Fig. 2a–e. Periodic-contrast lines appear with alternating spacings of 22 and  $8\text{ }\mu\text{m}$  in Fig. 2a which was taken with a film positioned  $\sim 5\text{ mm}$  away from the sample. They are images of the periodic inversion-domain walls. The vertical black-white lines or dots are images of segments or outcrops of dislocations. As the film-to-sample distance increases, the intensity variations become stronger, as shown in Fig. 2b–e. Furthermore, extra fringes appear in the domain-inverted regions of Fig. 2b–e compared with those of Fig. 2a. Clearly, the intensity variations observed above cannot be explained simply by a diffraction imaging or topographic effect. Instead, they mainly arise from interference, that is, the fringe patterns observed in Fig. 2b–e arise from phase contrast. As shown in Fig. 2a and e, the interference fringes run parallel to the domain

walls, but appear in the domain-inverted regions only. No interference fringes have been observed outside the treated regions except in the areas close to the border between the treated and untreated blocks. This indicates that the interference fringes are directly related to the presence of domain-inversion.

Figure 3a, b show the rocking curves of the 006 and 0012 reflections obtained using a Philips high-resolution Materials Research Diffractometer<sup>12</sup> in Warwick University. The measurements clearly indicate that the 006 rocking curve has been broadened significantly by domain-inversion processing, whereas the full-width at half-maximum (FWHM) of the 0012 reflection has not been much changed. Furthermore, the experimental FWHM of the 0012 reflection is much narrower than that of the 006 reflection in the domain-inverted regions. This is contrary to the result obtained in the untreated region where the experimental FWHM of the 0012 reflection is 1.7 times that of 006 because of the strong broadening effect of the instrument function on the high Bragg-angle reflection<sup>10</sup>. A further measurement via mapping is shown in

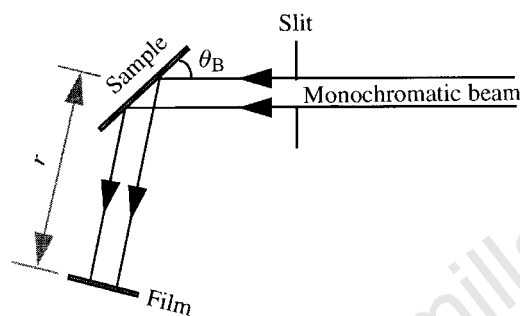
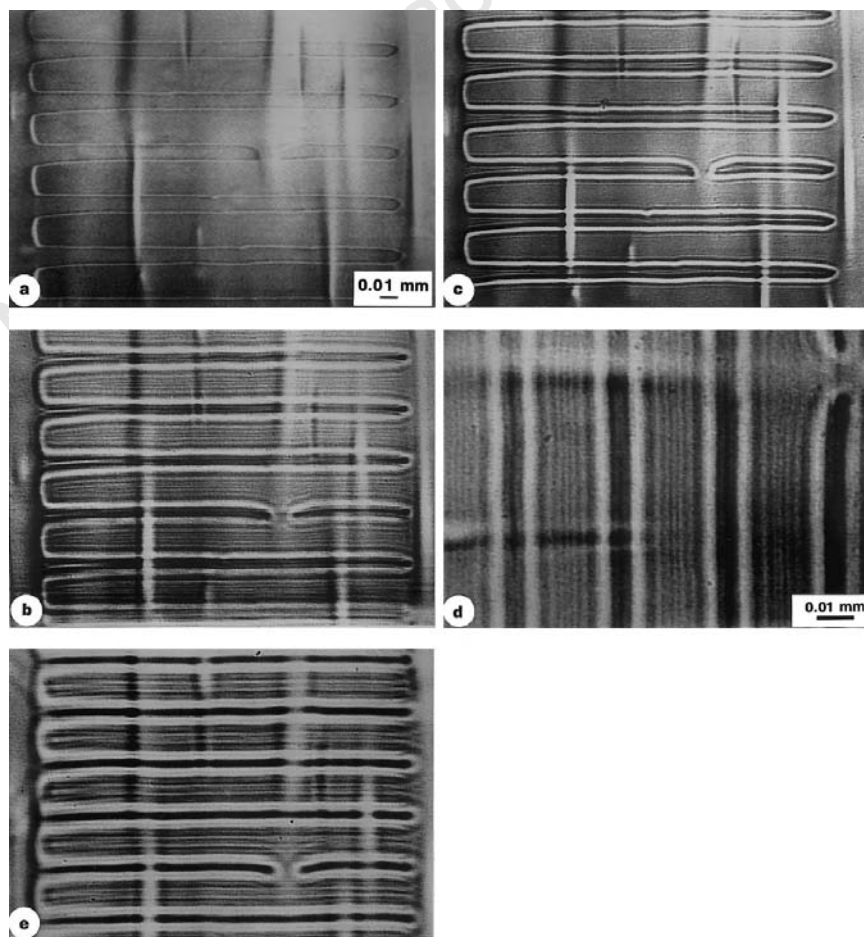


Fig. 3c and d, which shows two peaks overlapped along the  $\omega$ -scan direction in the 006 diffraction map in contrast to one single peak in the 0012 diffraction map. We note that the 006 symmetric reflection is much more sensitive to surface strains than the 0012 reflection because of the diffraction geometry and lower penetration depth for the former reflection<sup>10</sup>. Then, it is clear that the abnormal broadening of the 006 rocking curve results from the component of lattice tilts  $\delta\omega$  with respect to the goniometer axis in the top few micrometres of the surface between the inverted domains and original lattices.

It is the lattice tilting that is responsible for the spatial splitting and recombination of the coherently-related beams diffracted from the different parts for yielding phase contrast. This is analogous to the case of 'division of wavefront' (ref. 1) in optics where a monochromatic beam is divided by passage through a system such as apertures placed side by side. In our synchrotron experiment, the incident beam approximates to a linearly polarized monochromatic plane wave, and the LiNbO<sub>3</sub> sample should be

**Figure 1** Schematic drawing of the experimental configuration for taking X-ray phase-contrast images. A wavelength of 0.1 nm is selected by a perfect double-crystal Si (111) monochromator. The sample is set such that the domain walls are parallel to the plane of incidence of the X-ray beam. The vertical divergence of the incident synchrotron beam determined by the intrinsic width of the Si (111) reflection is  $\sim 2 \times 10^{-5}$  rad. The wavelength dispersion,  $\Delta\lambda/\lambda$ , of the output beam from the monochromator was  $\sim 10^{-4}$ , where  $\lambda$  is the X-ray wavelength and  $\Delta\lambda$  is the distribution of wavelengths. This resulted in a longitudinal coherence length,  $\delta^2/2\Delta\lambda$ , of about 0.5  $\mu\text{m}$ . The sample-to-source distance,  $R$ , was 40 m, the source size,  $d$ , was  $\sim 100 \mu\text{m}$ , the transverse coherence length,  $\lambda R/2d$ , was thus  $\sim 20 \mu\text{m}$  for the wavelength of 0.1 nm. Kodak high-resolution films were used to record the X-ray images.



**Figure 2** High-resolution phase-contrast images of a periodically domain-inverted LiNbO<sub>3</sub> crystal taken using the symmetric Bragg 006 reflection,  $2\theta_B = 25^\circ$ . The film-to-sample distances are 0.005 m (a), 0.1 m (b), 0.2 m (c), 0.2 m (d) and 0.5 m (e). The periodic black-white contrast with a periodicity of 30  $\mu\text{m}$  shown in a corresponds to the inversion-domain walls. The image in a is mainly of a conventional X-ray topographic nature, whereas intensity variations produced by interference are dominant in b-e where the interference occurs in periodically domain-inverted regions. Magnification in a, b, c and e is the same; d is an enlargement of a region in c, rotated anticlockwise by  $90^\circ$ .

treated as semi-infinitely thick considering the short extinction distance of  $4.4\text{ }\mu\text{m}$  for reflection 006 with wavelength  $0.1\text{ nm}$ . Therefore, there would be only one tie point or wave point (say,  $T_2$ ) excited on the dispersion surface<sup>13–15</sup> (Fig. 4a) and no interference would happen, unlike the Laue case if the sample was a perfect single crystal. However, because of the tilts of the lattice planes, the coherently-related beams diffracted essentially from two different misorientated parts are split inside the crystal, then recombined outside the crystal, which gives rise to interference. That is, a pair of tie points belonging to the same branch of the dispersion surface (say,  $T_2$  and  $T_2'$ ) are excited (Fig. 4b); these two tie points are directly related to the two misorientated parts tilted by an angle of  $\Delta\omega$  about the horizontal axis (although we discuss here the case of two split diffracted beams, this does not lead to a loss of generality). As the tie points  $T_2$  and  $T_2'$  are excited, the amplitudes at the sample surface,  $Z = 0$ , are given by

$$\mathbf{D}_o^i = \mathbf{D}_{o_2} + \mathbf{D}_{o_{2'}} \quad (1)$$

$$\mathbf{D}_g^e = \mathbf{D}_{g_2} + \mathbf{D}_{g_{2'}} \quad (2)$$

where  $\mathbf{D}_o^i$  and  $\mathbf{D}_g^e$  are the vector amplitudes of the outside incident and diffracted waves, respectively, and the subscripts 2 and 2'

represent the tie points  $T_2$  and  $T_2'$ . The outside diffracted waves from equation (2) can be written

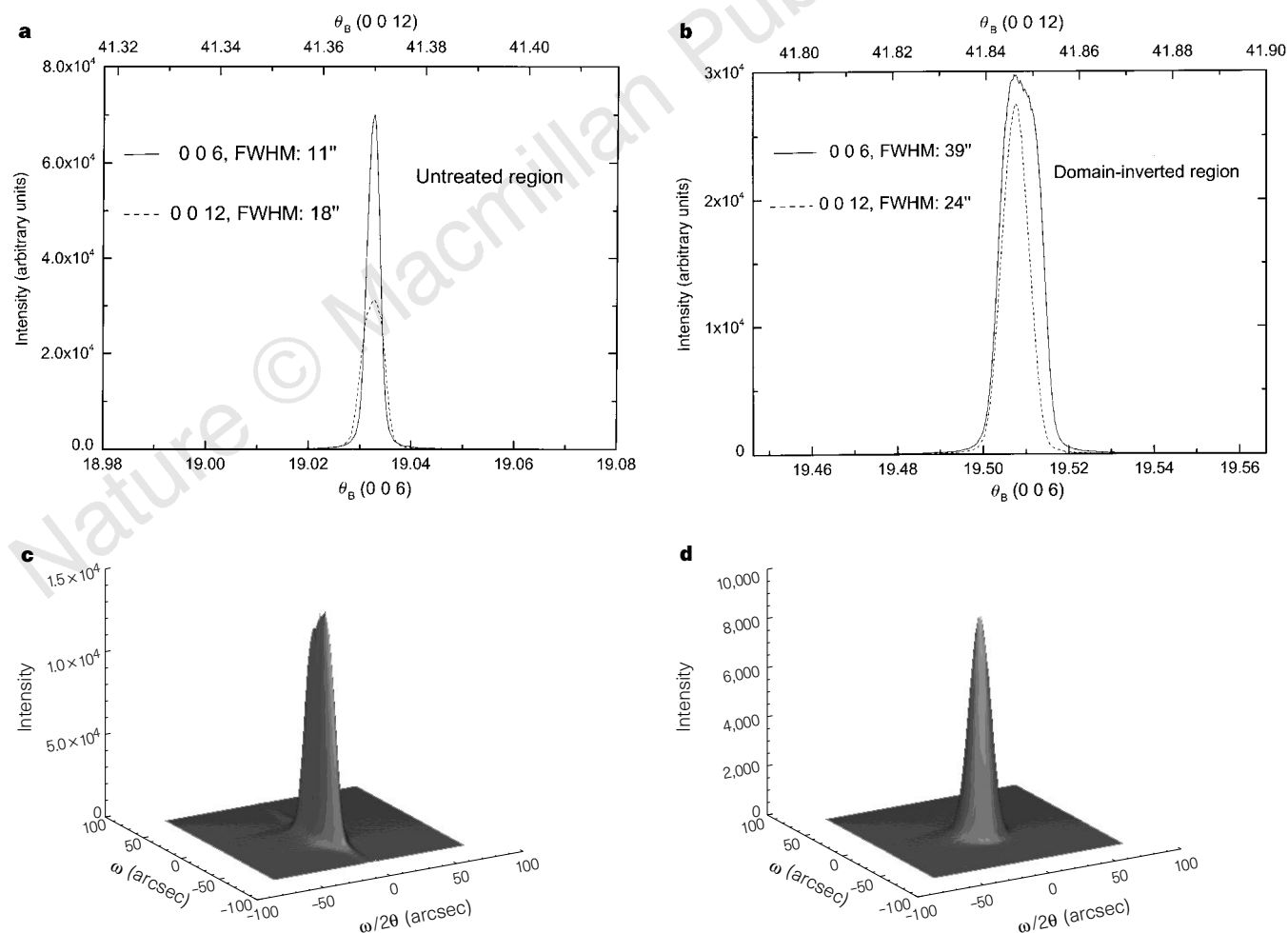
$$\mathbf{D}_g^e = D_{g_2} \exp[2\pi i(vt - \mathbf{K}_{g_2} \cdot \mathbf{r})] + D_{g_{2'}} \exp[2\pi i(vt - \mathbf{K}_{g_{2'}} \cdot \mathbf{r})] \quad (3)$$

where  $\mathbf{K}_{g_2}$  and  $\mathbf{K}_{g_{2'}}$  are the outside wavevectors of the diffracted waves excited for the tie points  $T_2$  and  $T_2'$ . These two wavevectors make a small angle of  $\Delta\phi$  ( $\Delta\phi = \Delta\omega$ ), and are different in magnitude by  $|\mathbf{K}_{g_{2'}} - \mathbf{K}_{g_2}|$  which determines the periodicity of interference fringes. The total diffracted intensity is

$$I = D_{g_2}^2 [1 + R^2 + 2R \cos 2\pi(\mathbf{K}_{g_{2'}} - \mathbf{K}_{g_2}) \cdot \mathbf{r}] \quad (4)$$

where  $R = D_{g_{2'}}/D_{g_2}$ , the ratio of the amplitudes of the diffracted waves of tie points  $T_2$  to  $T_2'$ .

The intensity given by equation (4) contains the phase factor,  $(\mathbf{K}_{g_{2'}} - \mathbf{K}_{g_2}) \cdot \mathbf{r}$ , and is modulated with a spatial periodicity  $|\mathbf{K}_{g_{2'}} - \mathbf{K}_{g_2}|^{-1}$ . The phase factor is independent of the choice of origin; the resultant pattern of interference actually represents a phase map of the periodically domain-inverted structure. The quantity  $|\mathbf{K}_{g_{2'}} - \mathbf{K}_{g_2}|^{-1}$  is equal to  $\lambda/\Delta\phi$ , where  $\lambda$  is the wavelength in vacuum (we note that the interference effect takes place outside



**Figure 3** High-resolution diffraction patterns and diffraction-space maps. **a** and **b** show a comparison between the 006 and 0012 rocking curves from the untreated region and those from the domain-inverted region. The FWHM of the 006 reflection is larger than that of the 0012 reflection measured from the domain-inverted region, which is contrary to the measurements obtained from the untreated region. **c** and **d** are the 006 and 0012 diffraction-space maps measured from the same domain-inverted region as that for the above rocking

curves, which shows that lattice tilting near the surface is responsible for the broadening of the 006 rocking curve. We note that the intensity distribution is much wider along the  $\omega$  than  $\omega/2\theta$  scan directions in the 006 diffraction-space map because of the overlapping of the two peaks. The sample is orientated so that the domain walls are perpendicular to the plane of incidence of the X-ray beam.



the crystal). Therefore, the periodicity of interference fringes is inversely proportional to the degree of lattice tilting, which is similar to the formation of Moiré fringes<sup>5,16–18</sup>. The tilts of the lattice planes vary with depth, from a few to 10 arcsec. The fringe patterns obtained at different sample-to-film distances arise from interference of different parts of the divergent cones defining the coherently-related beams. These different parts of the beams make different angles with each other, for example,  $\sim 5 \times 10^{-5}$ ,  $3 \times 10^{-5}$  and  $2 \times 10^{-5}$  rad in Fig. 2b, c, d, respectively. Clearly, it is impossible for the  $\mathbf{K}_{g_2}$  and  $\mathbf{K}_{g_2'}$  wavefronts with an angle of the order of  $10^{-5}$  rad to overlap inside the crystal or at the sample surface. These two diffracted wavefronts with a lateral displacement within the transverse coherence length of, say, 5–10  $\mu\text{m}$ , overlap at a distance of 10–100 mm from the sample, which yields interference. The phase-contrast nature of X-ray images would disappear if the detector-to-sample distance were not sufficiently large to ensure the overlap of the two coherently-related beams. The visibility of interference fringes, defined by  $V = (I_{\max} - I_{\min}) / (I_{\max} + I_{\min})$ , is equal to  $2R / (1 + R^2)$ , where  $I_{\max} = D_{g_2} (1 + R^2 + 2R)$  and  $I_{\min} = D_{g_2} (1 + R^2 - 2R)$ . The closer to the value of unity the ratio of the

amplitudes  $D_{g_2} / D_{g_2'}$ , the larger  $V$  is and the clearer the fringes are. As  $D_{g_2} / D_{g_2'} = 1$ ,  $I_{\max} = 4D_{g_2}^2$ ,  $I_{\min} = 0$ , the visibility of the fringes reaches 100%. Because of constructive and destructive interference the black and white contrast in Fig. 2b–d is very striking compared with the diffraction-dominant contrast in Fig. 2a. As expected from equation (4), the fringes run parallel to the horizontal direction. The fringe patterns revealed represent the lattice deformations in the domain-inverted regions. It appears that the strongest contrast occurs at the domain walls where the lattice deformation, and therefore the phase gradient, is different from the other regions. The interference fringes bend wherever both a tilt and a strain are present. The vertical white and black contrasts in Fig. 2a, b, c, e are the images of dislocations in which the horizontal fringes have been modified or even disrupted because of the presence of strong strain fields.

The exploitation of phase-contrast imaging using highly coherent X-ray beams provided by advanced synchrotron radiation sources opens up many possibilities for future research in materials science, condensed-matter physics and biological science. The present method uses the beams diffracted from the sample in either Bragg or Laue geometry for phase-contrast imaging, which differs from the Bonse–Hart interferometer-based technique<sup>5,19,20</sup> and the transmission-beam-based methods<sup>3,4</sup> and, therefore, allows non-destructive phase mapping of both highly and weakly absorbing single crystals. Because of the high sensitivity to deformation intrinsic to this method, it can be effectively used for the investigation of defects and microstructure in extensive bulk and thin-film materials such as used in optoelectronics and/or photonics. Use of the phase information obtained from imaging is expected to advance our understanding of the various post-growth modifications which have been extensively used in the fabrication of novel devices.

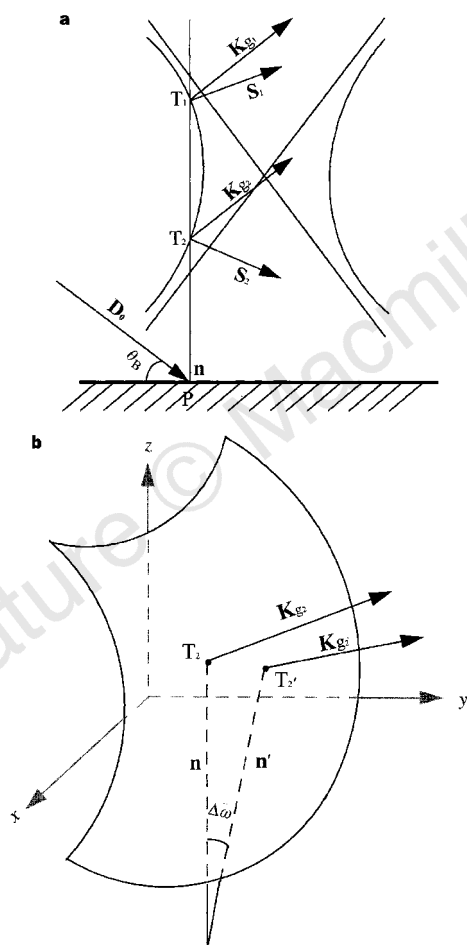
**Note added in proof:** It has recently come to our attention that there is another physical interpretation of the phase-contrast images shown here by P. Cloetens, P. Rejmankova and J. Baruchel (personal communication) which supposes a different physical origin for the phase jumps occurring at the domain walls and does not invoke lattice distortions. □

Received 1 May; accepted 4 November 1997.

1. Born, M. & Wolf, E. *Principles of Optics* 6th edn (Pergamon, Oxford, 1980).
2. Cowley, J. M. *Diffraction Physics* 3rd edn (North-Holland, Amsterdam, 1995).
3. Davis, T. J., Gao, D., Gureyev, T. E., Stevenson, A. W. & Wilkins, S. W. Phase-contrast imaging of weakly absorbing materials using hard x-rays. *Nature* **373**, 595–598 (1995).
4. Snigirev, A., Snigireva, I., Kohn, V., Kuznetsov, S. & Schelokov, I. On the possibilities of x-ray phase contrast microimaging by coherent high-energy synchrotron radiation. *Rev. Sci. Instrum.* **66**, 5486–5492 (1995).
5. Bonse, U. & Hart, M. An x-ray interferometer. *Appl. Phys. Lett.* **6**, 155–156 (1965).
6. Cloetens, P., Barret, R., Baruchel, J., Guigay, J.-P. & Schlenker, M. Phase objects in synchrotron radiation hard x-ray imaging. *J. Phys. D* **29**, 133–146 (1996).
7. Baruchel, J. in *X-ray and Neutron Dynamical Diffraction: Theory and Applications* (eds Authier, A., Lagomarsino, S. & Tanner, B. K.) 199–210 (Plenum, New York, 1996).
8. Armstrong, J. A., Bloembergen, N., Ducuing, J. & Pershan, P. S. Interactions between light waves in a nonlinear dielectric. *Phys. Rev.* **127**, 1918–1939 (1962).
9. Hu, Z. W., Thomas, P. A. & Webjörn, J. High-resolution x-ray characterization of periodically domain-inverted nonlinear optical crystals. *J. Phys. D* **28**, A189–A194 (1995).
10. Hu, Z. W., Thomas, P. A. & Webjörn, J. Observation of periodic domain inversion in periodically poled LiNbO<sub>3</sub> by high-resolution topography. *J. Appl. Crystallogr.* **29**, 279–284 (1996).
11. Hu, Z. W., Thomas, P. A. & Risk, W. P. Periodic domain inversion in poled KTiOPO<sub>4</sub> via high-resolution x-ray topography and diffraction-space mapping. *J. Phys. D* **29**, 2696–2704 (1996).
12. Fewster, P. F. A high-resolution multiple-crystal multiple-reflection diffractometer. *J. Appl. Crystallogr.* **22**, 64–69 (1989).
13. Authier, A. Trajet des rayons X dans un cristal parfait au voisinage de la réflexion totale. *J. Phys. Radium* **23**, 961–969 (1962).
14. Batterman, B. W. & Cole, H. Dynamical diffraction of x-rays by perfect crystals. *Rev. Mod. Phys.* **36**, 681–716 (1964).
15. Tanner, B. K. *X-ray Diffraction Topography* (Pergamon, Oxford, 1976).
16. Lang, A. R. & Miuscov, V. F. Å-scale displacements revealed by x-ray Moiré topographs. *Appl. Phys. Lett.* **7**, 214–216 (1965).
17. Bonse, U. & Hart, M. Moiré patterns of atomic planes obtained by x-ray interferometry. *Z. Phys.* **190**, 455–467 (1966).
18. Lang, A. R. X-ray Moiré topography of lattice defects in quartz. *Nature* **220**, 652–657 (1968).
19. Hart, M. Bragg reflection x-ray optics. *Rep. Prog. Phys.* **34**, 435–490 (1971).
20. Ando, M. & Hosoya, S. in *Proc. 6th Int. Conf. on X-ray Optics and Microanalysis* (eds Shinoda, G. et al.) 63–68 (Univ. Tokyo Press, 1972).

**Acknowledgements.** Z.W.H. was on leave from the National Laboratory of Solid State Microstructures, Nanjing University, China. Z.W.H. thanks A. R. Lang for discussions, and was supported by the UK EPSRC.

Correspondence and requests for materials should be addressed to P.A.T. (e-mail: phrve@csv.warwick.ac.uk).



**Figure 4** Dispersion surface showing the tie points of wavefields in the symmetric Bragg case. **a**, Two-dimensional drawing of the dispersion surface in the normal case. The tie points,  $T_1$  and  $T_2$  on dispersion surface (hyperbola) are determined by the entrance point  $P$  and the crystal surface normal  $\mathbf{n}$ . Only the tie point  $T_2$ , which has its Poynting vector pointing into the crystal, is excited for a semi-infinitely thick crystal. **b**, 'Three-dimensional' drawing of a segment of the dispersion surface to show the division of the wavefields by the lattice tilting between two parts of a crystal. Because of lattice tilting, the crystal surface normal is split into  $\mathbf{n}$  and  $\mathbf{n}'$ , which leads to the excitation of two wave points,  $T_2$  and  $T_2'$ . Interference would take place between the wavefield of wavevector  $\mathbf{K}_{g_2}$  and that of wavevector  $\mathbf{K}_{g_2'}$ . Note that the wave points,  $T_2$  and  $T_2'$ , have their Poynting vectors pointing into the crystal.

# Identification of cathode materials for lithium batteries guided by first-principles calculations

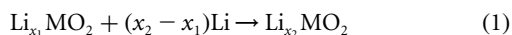
G. Ceder, Y.-M. Chiang, D. R. Sadoway, M. K. Aydinol, Y.-I. Jang & B. Huang

Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139-4307, USA

Lithium batteries have the highest energy density of all rechargeable batteries and are favoured in applications where low weight or small volume are desired—for example, laptop computers, cellular telephones and electric vehicles<sup>1</sup>. One of the limitations of present commercial lithium batteries is the high cost of the LiCoO<sub>2</sub> cathode material. Searches for a replacement material that, like LiCoO<sub>2</sub>, intercalates lithium ions reversibly have covered most of the known lithium/transition-metal oxides, but the number of possible mixtures of these<sup>2–5</sup> is almost limitless, making an empirical search labourious and expensive. Here we show that first-principles calculations can instead direct the search for possible cathode materials. Through such calculations we identify a large class of new candidate materials in which non-transition metals are substituted for transition metals. The replacement with non-transition metals is driven by the realization that oxygen, rather than transition-metal ions, function as the electron acceptor upon insertion of Li. For one such material, Li(Co,Al)O<sub>2</sub>, we predict and verify experimentally that aluminium substitution raises the cell voltage while decreasing both the density of the material and its cost.

The important characteristics of a lithium/metal-oxide compound for battery applications include the voltage at which it exchanges lithium, the amount of lithium that can be reversibly intercalated, and the stability of the material. The first two properties determine the energy density, and the latter limits the lifetime of the battery. Although a high voltage is desirable for obtaining high energy density<sup>3</sup>, the practical use of high-voltage cathodes is at present limited by the voltage range over which the electrolyte is stable. The recent development of high-voltage electrolytes<sup>6</sup> is, however, likely to ameliorate this problem.

We demonstrated recently<sup>7,8</sup> that the average potential for intercalation between two compositions  $x_1$  and  $x_2$  can be obtained from the energy change in the reaction:



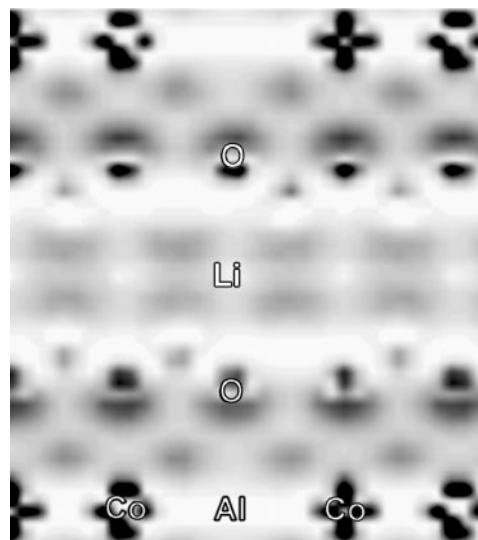
The advantage of a first-principles computation is that it can be used to calculate the energy of the three compounds in equation (1), and hence the average potential, for any metal M (or combination of them), in any structure, whether these conditions have been experimentally realized or not. To calculate the total energies in equation (1) we use the *ab initio* ultrasoft-pseudopotential method as implemented in ref. 9. For the reactions  $\text{Li} + \text{Mn}_2\text{O}_4 \rightarrow \text{LiMn}_2\text{O}_4$  and  $\text{Li} + \text{CoO}_2 \rightarrow \text{LiCoO}_2$  we find average potentials of 4.0 and 3.75 V, respectively. The measured values are approximately 4.1 and 4.0 V, indicating that we have reasonable agreement with experiment, even though there are no adjustable parameters in the theory. The errors are consistent with the systematic underprediction of the potential in these compounds by  $\sim 0.2$  V (ref. 7). Clearly, the *ab initio* prediction of intercalation voltages is possible. New materials can therefore be pre-screened before attempting their synthesis.

It is traditionally assumed that the intercalation voltage is determined by the redox potential of the transition-metal ion

which is believed to change valence on Li insertion. We recently argued that for the late-transition-metal oxides it is instead the oxygen which is in large part responsible for the electron exchange<sup>7</sup>. As electron exchange with oxygen leads to a lower electrostatic energy (hence a higher voltage), we speculated that increased voltage correlates with increased oxygen participation in the electron exchange. If this is correct, part or all of the transition metal in a cathode material could be replaced by non-transition-metal ions, while retaining electrochemical Li-activity, although at a higher voltage. Such substitution would significantly increase the number of compounds that can be considered as cathode-active materials. In particular, elements from the third row of the periodic table would be of interest as they have lower mass than the 3d metals.

In this work, we first calculated the hypothetical intercalation voltage in Li<sub>x</sub>AlO<sub>2</sub> and found it to be 5.4 V. LiAlO<sub>2</sub> has empty Al *p* orbitals well above the filled oxygen *p* states with no *d* states in between (as have all the transition-metal oxides) and electron exchange therefore occurs completely with the oxygen bands. Experimentally, this Li intercalation capacity will probably not be accessible as LiAlO<sub>2</sub> is electronically insulating. Electronic conductivity is required in an intercalation oxide, as Li<sup>+</sup> diffusion can occur only when accompanying electron motion is also possible. However, in solid solution with other transition-metal oxides, LiAlO<sub>2</sub> may be an attractive component. If our prediction regarding electrochemical activity of the oxygen ion is correct, Al substitution for transition-metal oxides will lead to higher Li intercalation voltage, as the fixed 3+ valence of Al forces more electron exchange with oxygen. However, our calculations indicate potential difficulties for mixing LiAlO<sub>2</sub> with other lithium/metal-oxides. We calculated the formation enthalpy for Li(M,Al)O<sub>2</sub> mixtures with M = Ti, V, Fe, Co and found them all to be positive, indicating that, at least at low temperature, separate regions of LiAlO<sub>2</sub> and LiMO<sub>2</sub> are favoured over the solid solution. Only in Li(Co,Al)O<sub>2</sub> is the formation enthalpy small enough (30 meV) to allow for entropy-driven mixing. We note that previous investigations of Al-doping of transition-metal oxides show no apparent effect on the intercalation potential<sup>10,11</sup>. However, phase separation at a scale not detectable by X-rays would annihilate the effect of Al doping on the potential, and result in a material with properties indistinguishable from the undoped material.

For Li(Al<sub>0.33</sub>Co<sub>0.67</sub>)O<sub>2</sub> and Li(Al<sub>0.67</sub>Co<sub>0.33</sub>)O<sub>2</sub> in the  $\alpha$ -NaFeO<sub>2</sub> structure, our first-principles calculations predict an average Li intercalation potential of respectively 4.2 and 4.7 V. Given the



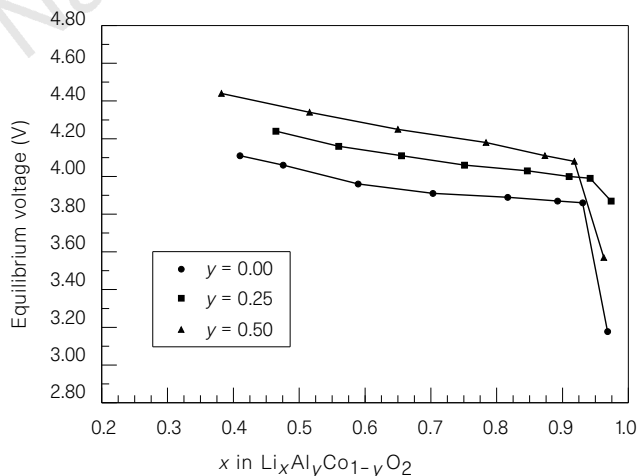
**Figure 1** Positive part of the electron density difference between Li(Al<sub>0.33</sub>Co<sub>0.67</sub>)O<sub>2</sub> and (Al<sub>0.33</sub>Co<sub>0.67</sub>)O<sub>2</sub> in a plane perpendicular to the direction of layering in the structure. Darker indicates larger electron density.

consistent underestimation of experimental voltages by the calculations<sup>7</sup>, the true potential is probably  $\sim 0.2$  V higher. The increase in voltage over pure  $\text{LiCoO}_2$  is in agreement with our understanding of the increasing role of oxygen in Al-doped material. Figure 1 shows the change in charge density of the valence electrons when going from  $(\text{Al}_{0.33}\text{Co}_{0.67})\text{O}_2$  to  $\text{Li}(\text{Al}_{0.33}\text{Co}_{0.67})\text{O}_2$ . These results are obtained by subtracting the charge density in  $(\text{Al}_{0.33}\text{Co}_{0.67})\text{O}_2$  from that in  $\text{Li}(\text{Al}_{0.33}\text{Co}_{0.67})\text{O}_2$ . For these calculations, all atomic positions were assumed to remain unchanged on lithiation (whereas all coordinates were fully relaxed for the voltage calculations). Although Co accepts some charge from the extra lithium atom, none is transferred to Al. Considerable charge transfer to oxygen is clearly visible.

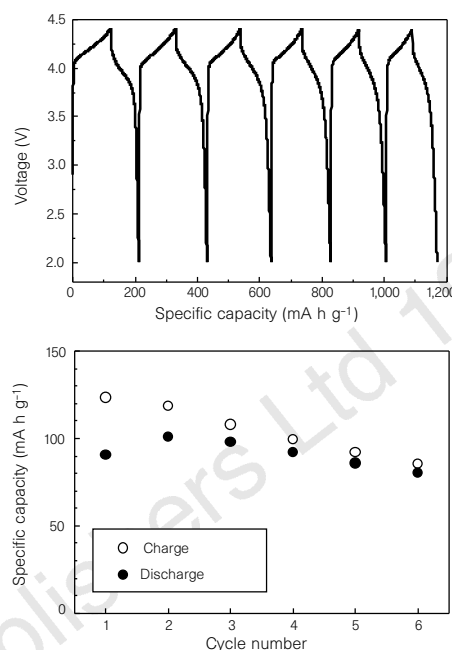
To test the predictions of the computations we developed a method to synthesize homogeneous  $\text{Li}(\text{Co,Al})\text{O}_2$  solid solutions in the  $\alpha\text{-NaFeO}_2$  structure. First, mixed cobalt–aluminium hydroxides were precipitated from mixed aqueous solutions containing  $\text{Co}(\text{NO}_3)_2$  and  $\text{Al}(\text{NO}_3)_3$ , added dropwise to a continuously stirred  $\text{LiOH}$  aqueous solution kept at pH 10.5. Nitrate ions were removed from the precipitate through a rinsing procedure described in ref. 12. Lithium was then added through a freeze-drying procedure, in which the mixed hydroxide was dispersed in aqueous  $\text{LiOH}$  solution yielding a total  $\text{Li}/(\text{Co} + \text{Al})$  molar ratio of unity, the suspension was atomized into liquid nitrogen, and the frozen droplets were freeze-dried.

Transmission electron microscopy showed this precursor to consist of crystalline Co,Al hydroxide uniformly mixed with amorphous Li hydroxide at the submicrometre scale. X-ray powder diffraction showed that a single-phase solid solution in the  $\alpha\text{-NaFeO}_2$  structure (space group  $R\bar{3}m$ ) is obtained on firing this precursor to  $T > 400^\circ\text{C}$  in air for the compositions  $y = 0.25$  and  $y = 0.5$ . For  $y = 0.75$ , a two-phase mixture of the  $\alpha\text{-NaFeO}_2$  solid solution and the tetragonal phase of  $\text{LiAlO}_2$  was obtained.

Cathodes containing  $\text{LiAl}_y\text{Co}_{1-y}\text{O}_2$  were prepared by mixing the oxide powder with carbon black and potting the mixture in a binder of poly(vinylidene fluoride). Pellets weighing 10–35 mg and measuring  $1.0\text{ cm}^2$  in surface area were formed by pressing at 360 MPa. These were dried at  $140^\circ\text{C}$  under primary vacuum for 24 h and transferred to an argon-filled glove box. The test cell consisted of two stainless-steel electrodes in a holder made of Teflon. The anode was a lithium ribbon 0.75 mm thick. The electrolyte consisted of a 1 M solution of  $\text{LiPF}_6$  in ethylene carbonate–diethyl carbonate, 1:1 by volume. The separator was a film of Celgard 2400 (Celgard LLC,



**Figure 2** Open-circuit potential as a function of lithium content in pure  $\text{LiCoO}_2$  and in Al-doped materials.



**Figure 3** Cycling behaviour of  $\text{Li}(\text{Al}_{0.25}\text{Co}_{0.75})\text{O}_2/\text{Li}$  and specific capacity between 2.0 and 4.4 V. The current density was  $0.4\text{ mA cm}^{-2}$ .

Charlotte, North Carolina). Charge/discharge tests were performed with a Maccor Series 4000 Automated Test System (Maccor Inc., Tulsa, Oklahoma).

Figure 2 shows the open-circuit voltage (equilibrium potential) as a function of Li content in the oxide for samples calcined at  $850^\circ\text{C}$ . These results were obtained by charging the cell in steps and allowing it to equilibrate for 15 h after each step. The equilibrium potential of the cells increases systematically with the Al content in the oxide, confirming the theoretical predictions. We note that at a value of  $1 - x = 0.4$  the equilibrium voltage of the  $\text{Li}_{1-x}\text{Al}_{0.5}\text{Co}_{0.5}\text{O}_2/\text{Li}$  cell was as high as 4.40 V. As the calculations give the average voltage between compositions  $\text{MO}_2$  and  $\text{LiMO}_2$ , whereas the cells were charged over only part of this composition range, direct quantitative comparison of the theoretical and experimental voltage is not yet possible.

Figure 3 shows the first few charges and discharges between 2.0 and 4.4 V (current density of  $0.4\text{ mA cm}^{-2}$ ) for the  $\text{LiAl}_{0.25}\text{Co}_{0.75}\text{O}_2/\text{Li}$  cell. The initial charge and discharge capacity were 120 and  $90\text{ mA h g}^{-1}$ , respectively. On the sixth cycle, charge and discharge in the same potential window resulted in a capacity of  $\sim 90$  and  $80\text{ mA h g}^{-1}$ , respectively. The 4.4 V charging limit is imposed by the stability of the electrolyte used in these experiments and it is likely that more capacity could be found at higher potentials. The material clearly shows some capacity fade, as has been the case with many Li insertion materials in the initial stage of their development.

We expect similar results to these in other  $\text{LiAl}_y\text{M}_{1-y}\text{O}_2$  solid solutions, where M is a transition metal. The present results show that the intercalation potential of such solid solutions, and probably those of other structures such as the spinel, can be 'quantum-engineered'. Owing to its low weight and price, aluminium can significantly increase energy density while reducing cost. Commercial feasibility of any new material will, however, also rely on successful synthesis and demonstration of high capacity and stability under electrochemical cycling. □

Received 20 August 1997; accepted 9 February 1998.

1. Pistoia, G. *Lithium Batteries* (Elsevier, Amsterdam, 1994).
2. Delmas, C. & Saadoun, I. Electrochemical and physical properties of the  $\text{Li}_x\text{Ni}_{1-y}\text{Co}_y\text{O}_2$  phases. *Solid State Ionics* **53–56**, 370–375 (1992).
3. Fey, G., Li, W. & Dahn, J. R.  $\text{LiNiVO}_4$ : A 4.8 Volt electrode material for lithium cells. *J. Electrochem. Soc.* **141**, 2279–2282 (1994).
4. Reimers, J. N., Rosen, E., Jones, C. D. & Dahn, J. R. Structure and electrochemistry of  $\text{Li}_x\text{Fe}_{1-y}\text{Ni}_y\text{O}_2$ . *Solid*



State Ionics 61, 335–334 (1993).

5. Ohzuku, T., Ueda, A., Nagayama, M., Iwakoshi, Y. & Komori, H. Comparative study of  $\text{LiCoO}_2$ ,  $\text{LiNi}_{1/2}\text{Co}_{1/2}\text{O}_2$  and  $\text{LiNiO}_2$  for 4 volt secondary lithium cells. *Electrochim. Acta* 38, 1159–1167 (1993).
6. Tarascon, J. M. & Guyomard, D. New electrolyte compositions stable over the 0 to 5 V voltage range and compatible with the  $\text{Li}_{1+x}\text{Mn}_2\text{O}_4/\text{carbon}$  Li-ion cells. *Solid State Ionics* 69, 293–305 (1994).
7. Aydinol, M. K., Kohan, A. F., Ceder, G., Cho, K. & Joannopoulos, J. Ab-initio study of lithium-intercalation in metal-oxides and metal-dichalcogenides. *Phys. Rev. B* 56, 1354–1365 (1997).
8. Ceder, G., Aydinol, M. K. & Kohan, A. F. Application of first-principles calculations to the design of rechargeable Li-batteries. *Comput. Mater. Sci.* 8, 161–169 (1996).
9. Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* 6, 15–50 (1996).
10. Nazri, G. A., Rougier, A. & Kia, K. F. In *Solid State Chemistry of Inorganic Materials* (eds Davies, P. K., Jacobson, A. J., Torardi, C. C. and Vanderah, T. A.) 635–646 (Mater. Res. Soc. Symp. Proc. 453, Materials Research Society, Pittsburgh, 1997).
11. Ohzuku, T., Ueda, A. & Kouguchi, M. Synthesis and characterization of  $\text{LiAl}_{1/4}\text{Ni}_{3/4}\text{O}_2(\text{R3m})$  for lithium-ion (shuttlecock) batteries. *J. Electrochem. Soc.* 142, 4033–4039 (1995).
12. Chiang, Y.-M. et al. Synthesis of  $\text{LiCoO}_2$  by decomposition and intercalation of hydroxides. *J. Electrochem. Soc.* 145, 887–891 (1998).

**Acknowledgements.** We thank A. Mayes for discussions and comments. This work was sponsored by Furukawa Electric and the US Department of Energy through the Idaho National Engineering Laboratory University Research Consortium.

Correspondence and requests for materials should be addressed to G.C. (e-mail: gerd@lanai-mit.edu).

## Surface topography dependence of biomolecular hydrophobic hydration

Yuen-Kit Cheng & Peter J. Rossky

Department of Chemistry and Biochemistry, University of Texas at Austin, Austin, Texas 78712-1167, USA

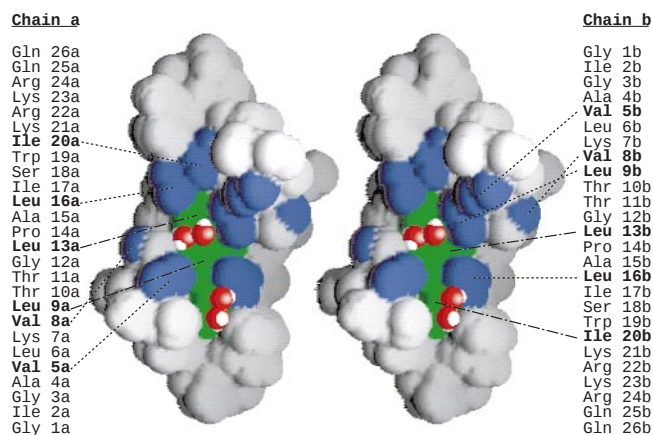
Many biomolecules are characterized by surfaces containing extended nonpolar regions<sup>1</sup>, and the aggregation and subsequent removal of such surfaces from water is believed to play a critical role in the biomolecular assembly in cells<sup>2</sup>. A better understanding of the hydrophobic hydration of biomolecules may therefore yield new insights into intracellular assembly. Conventional views hold that the hydration shell of small hydrophobic solutes is clathrate-like, characterized by local cage-like hydrogen-bonding structures and a distinct loss in entropy<sup>2</sup>. The hydration of extended nonpolar planar surfaces, however, appears to involve structures that are orientationally inverted relative to clathrate-like hydration shells<sup>3,4</sup>, with unsatisfied hydrogen bonds that are directed towards the hydrophobic surface. Here we present computer simulations of the interaction between the polypeptide melittin and water that demonstrate that the two different hydration structures also exist near a biomolecular surface. We find that the two structures are distinguished by a substantial difference in the water–water interaction enthalpy, and that their relative contributions depend strongly on the surface topography of the melittin molecule: clathrate-like structures dominate near convex surface patches, whereas the hydration shell near flat surfaces fluctuates between clathrate-like and less-ordered or inverted structures. The strong influence of surface topography on the structure and free energy of hydrophobic hydration is likely to hold in general, and will be particularly important for the many biomolecules whose surfaces contain convex patches, deep or shallow concave grooves and roughly planar areas<sup>5</sup>.

The magnitude of the free energy of hydrophobic hydration, as inferred from the solubilities of small hydrocarbons in water, is  $\sim 25 \text{ cal mol}^{-1} \text{ \AA}^{-2}$  and is typically assumed to depend predominantly on the solute surface area<sup>6</sup>. This, and similar values derived from the data<sup>6</sup>, are relatively low when compared with (for instance) the corresponding macroscopic measure of the liquid pentane–water interface, which is  $\sim 72 \text{ cal mol}^{-1} \text{ \AA}^{-2}$ . Although the origin of this large difference remains controversial (see ref. 7 for a review), it is reasonable to anticipate that the molecularity of the solvent and solute are relevant to the value obtained.

Amphipathic helical polypeptides (with segregation of hydrophilic and hydrophobic moieties) have been chosen as the target in this study. Melittin, a 26-mer polypeptide found in honeybee venom, is one of the most studied archetypes of membrane proteins<sup>8</sup>. It readily self-associates as a tetramer in an aqueous solution of high peptide concentration, high pH, or high ionic strength<sup>8</sup>. Crystallized as a tetramer, it is spatially a dimer of two almost stereochemically identical amphipathic  $\alpha$ -helical dimers related by a two-fold symmetry axis<sup>9</sup>. Each dimer possesses a relatively flat hydrophobic surface, surrounded by convex patches of hydrophobic moieties (Fig. 1); the flat surface is completely buried in the centre of the tetramer on tetramerization<sup>9</sup>. Self-association of  $\alpha$ -helical melittin monomers is thought to be important in its lytic activity of membranes, although the actual mechanism is still unknown. (See ref. 8 for a review.)

The interfacial hydration properties are expected to be crucial to the role of hydrophobic hydration of membrane proteins preceding and during membrane association. As the hydration of extended hydrophobic surfaces is apparently common in biological processes (for example, as inferred in the functioning surfaces of chaperonins<sup>10</sup>), results regarding hydrophobic hydration of surfaces, explored in this study, are of significance to a wider range of processes and compounds than the mechanism of melittin in membrane lysis.

In the following analysis, the surface of a melittin dimer is partitioned into two main types, ‘flat’ and ‘convex’, according to the local surface curvature (see Fig. 1). The coordinates of the dimer were extracted from the crystal structure of melittin tetramer<sup>9</sup> deposited in the Brookhaven Protein Data Bank<sup>11</sup> and its structure was kept rigid throughout the simulation. The use of a fixed protein structure is not implemented as a simplification, but rather is required if one is to specifically identify the nature of hydration that is compatible with a given geometry of the protein surface. It is of interest also to learn how this protein structure might relax in response to full solvent exposure, but the present approach is needed if one is to understand the physical reasons for that process. Studies on the dynamics and other aspects of monomeric melittin in solution can be found elsewhere<sup>12</sup>. The present simulation (see Methods section) was performed at a temperature of 300 K in the



**Figure 1** Stereographic image of the hydrophobic surface of the melittin dimer<sup>9</sup>; this is completely buried on tetramerization. The peptide sequences are provided in standard notation and are aligned with the model; the surface carbon centres selected for hydration analysis are indicated in bold. The central contiguous segment (shown green) is relatively flat and long (accessible surface area,  $72.4 \text{ \AA}^2$ ); the small convex surface patches are shown blue. Four surface proximal water molecules (rendered with oxygen radius  $r_{\text{O}} = 1.6 \text{ \AA}$  and hydrogen radius  $r_{\text{H}} = 1.0 \text{ \AA}$ ; oxygen atoms are shown in red) are overlaid on the images. Surface regions not considered in this analysis are shown white. The picture is rendered by the program GRASP<sup>22</sup>, and accessible surface area calculations by GEPOL93<sup>16</sup>.

microcanonical ensemble. Equilibration took  $\sim 23$  ps of model time; a further 120 ps of simulation was performed for subsequent analyses.

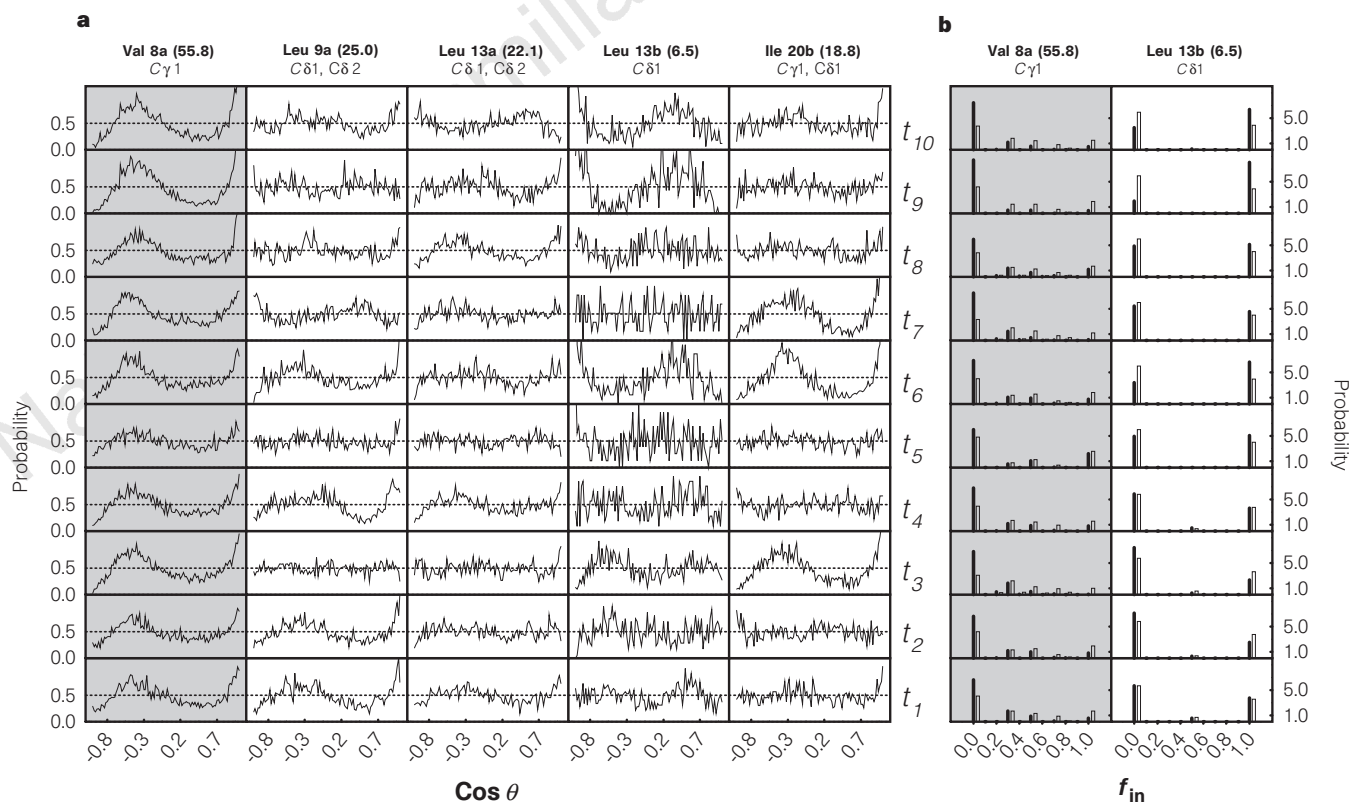
For analysis, we define the two OH bonds and the two 'lone pair' directions of each water molecule, pointing tetrahedrally outwards from the oxygen atom, as four hydrogen-bonding (hb) vectors. Further, water molecules which are closer to a particular solute atom than to any of the other solute atoms (and within 4.0 Å) are considered proximal to that solute atom<sup>13</sup>.

The orientation of each proximal water molecule with respect to the local surface normal can be characterized by  $\theta$ , which is the angle between each of the hb vectors and the outward radial direction pointing from the carbon nucleus associated with the surface towards the water oxygen atom. Figure 2a depicts the calculated distributions<sup>14</sup> of  $\cos \theta$ . Classical clathrate-like hydration is evidenced by a distribution<sup>14</sup> such as that observed for the hydration of residue Val 8a, which belongs to the convex region (see Fig. 1). This has a rather broad maximum around  $\cos \theta_t$  (where  $\theta_t$  is the tetrahedral angle) ( $\sim -0.336$ ) and a sharp rise close to  $\cos(0) = 1$ . This corresponds to three of the four hb vectors of each water molecule being orientated nearly tangentially to its proximal surface atom, and one vector pointing away. In contrast, inverted hydration of water molecules proximal to the surface would have a characteristic distribution which mirrors that of the clathrate-like one, maximizing at  $\cos \theta = -1$  and  $+0.336$ , with one hb vector pointing directly into the surface. In fact, we find that for the melittin interface, water molecules proximal to the residues belonging to the flat central region show orientational fluctuations (Fig. 2a) among

strongly clathrate-like distribution (for example, Ile 20b: time segments  $t_3$ ,  $t_6$  and  $t_7$ ), a weak clathrate-like form (for example, Leu 13a:  $t_4$ ), an inverted structure (for example, Leu 13b:  $t_9$  and  $t_{10}$ ), and a mixed behaviour (clathrate and inverted leading to a comparatively flat curve). As each frame in the figure represents a 12 ps average, it is evident that the prevalence of any one structural type typically persists for  $\sim 10$ –20 ps.

The hydration structure of the central hydrophobic region cannot be characterized by either a classic, small-molecule, clathrate-like structure<sup>14</sup> or the orientationally inverted structure characteristic of truly extended surfaces<sup>3</sup>, but rather shows a fluctuating distribution over both. To characterize the degree to which the structural fluctuations are collective contributions from the water molecules, we consider the quantity  $f_{in}$  (the ratio of the number of proximal water molecules with any one of its hb vectors pointing radially inward towards the solute atom (defined as  $\cos \theta \leq -0.8$ ), divided by the total number of proximal water molecules of this solute atom) in each configuration. As illustrated in Fig. 2b, our simulations indicate a high probability of clathrate-like structures ( $f_{in} = 0$ ) dominating near the convex surface patch of Val 8a, and that the probability of encountering inverted structures ( $f_{in} = 1$ ) increases markedly near the relatively flat region. This shift is particularly evident for residue Leu 13b during part of the trajectory.

These results clearly show a topography dependence for hydrophobic hydration of a molecular surface. To evaluate the energetic consequences, we compute the average binding energy,  $E_b$  (the interaction of a molecule with all other molecules in the system) of proximal water molecules in each surface set. Except for those



**Figure 2** Statistical distribution of the orientation of proximal water molecules in ten consecutive time segments ( $t_1$  to  $t_{10}$ ). Each curve/histogram represents a 12 ps normalized average. **a**, Statistical distribution of the orientation of proximal water molecules with respect to the local surface normal (as measured by  $\theta$ ; see text). Only results of proximal water molecules around indicated selected residues are shown. Convex patches are highlighted with a shaded background and represented by Val 8a; Val 5a, Leu 16a, Val 5b, Val 8b and Leu 16b are similar. The solvent-accessible surface areas (in Å<sup>2</sup>) of the indicated surface carbon atoms in each residue are shown in parentheses beside the residue labels.

Horizontal dotted lines at 0.5 represent the hypothetical result for random solvent orientation. **b**, Statistical distribution of the inward pointing fraction  $f_{in}$  of hb vectors for proximal water molecules (unfilled bars: analytical results for a hypothetical random orientation on the same surface and with the same number of proximal water molecules; filled bars: observed result of simulation). All plots are integrally normalized. Typical results around convex surface patches are represented by Val 8a (highlighted by shaded background), and the results around the flat surface by Leu 13b.

around residue Ile 20a, which is spatially close to one of the charged termini, water molecules proximal to the hydrophobic face of the melittin dimer have very small interactions with the counterions in the system, and each has a total average interaction with the solute protein of at most  $-1.0 \text{ kcal mol}^{-1}$ . Therefore, for further analysis, we consider only the binding energies resulting from the water–water interaction alone,  $E_{\text{bww}}$ . As illustrated in Fig. 3a, in the vicinity of convex patches such as residue Val 8a, the proximal water molecules have an average interaction energy  $E_{\text{bww}}$  slightly higher (less favourable) than that of the bulk molecules. In the flat region, however, the average interaction energies are significantly higher. Most extreme, in the proximity of residue Leu 13b,  $E_{\text{bww}}$  increases by  $\sim 5 \text{ kcal mol}^{-1}$  (corresponding to 25%) relative to the bulk value. A reasonable estimate of the number of hydrogen bonds associated with one water molecule,  $n_{\text{hb}}$ , can be obtained by determining the average number of water–water pair interactions that are associated with a given molecule and whose energies are less than or equal to  $-3 \text{ kcal mol}^{-1}$  (ref. 15). Figure 3b, which shows  $n_{\text{hb}}$  values for both proximal and bulk water molecules, corroborates the pattern seen for the interaction energies: water interactions in the bulk and around convex patches are comparable, whereas the extent of favourable interactions near flat surfaces is substantially reduced.

The flat group (Leu 9a, Leu 13a and residue Leu 13b) has a total solvent-accessible surface area<sup>16</sup> of  $53.6 \text{ \AA}^2$  and a total coordination number of proximal water molecules (results not shown) similar to those of the convex group (residue Val 8a). Nevertheless, the solvent orientation (Fig. 2a),  $E_{\text{bww}}$  (Fig. 3a) and  $n_{\text{hb}}$  (Fig. 3b) are significantly different. We conclude that, depending on the local curvature, hydrophobic residues of even the same chemical type can impose constraints on hydrogen bonding that lead to significantly different orientational structures and energetics of the hydration shell.

The flat surface can be solvated not only by a clathrate-like hydration shell, but also by an inverted one which is characterized by a less favourable water–water binding energy. The distinct fluctuations between clathrate-like structures and less ordered or inverted structures show that these structures are of comparable free energy. Thus the inverted structure must be entropically more stable than the clathrate-like structure. Although the present results show large energetic effects due to the two types of hydration, our findings do not directly address the issue of the free-energy change that would be associated with removal of the melittin interface from aqueous exposure, particularly as enthalpy–entropy compensation is expected. However, even in the extreme case of essentially complete enthalpy–entropy compensation at a given temperature, the temperature dependence of hydrophobic interaction associated with melittin hydration should be distinctly different from that associated with small-molecule hydration<sup>2,17</sup>, since the enthalpy determines this dependence.

A practical question relates to the required extent of this surface for inverted hydration structure to persist. Evidently, it need not be

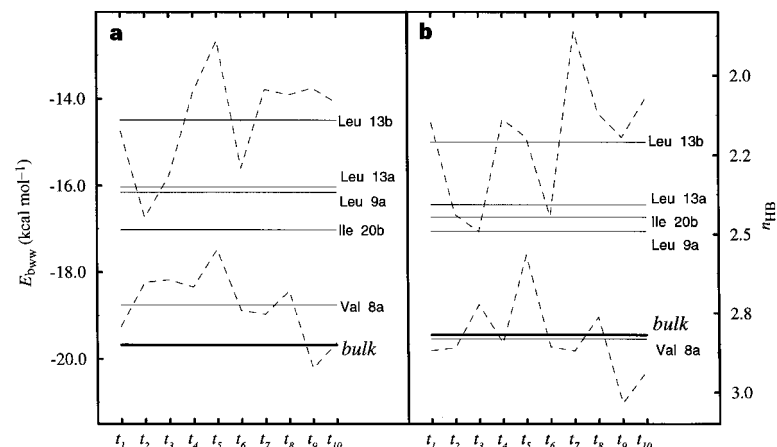
large—the flat surface considered here involves no more than seven water molecules in its first hydration shell. On the other hand, proximal water molecules around extended hydrophobic surfaces in the vicinity of polar and charged groups (for example, Tyr residues of the dimeric interface of insulin) are able to maintain a total  $E_{\text{b}}$  close to that of the bulk (Y.-K.C. and P.J.R., unpublished results). This observation suggests that both the degree of spatial dispersion of the hydrophobic elements of a biomolecular surface and the frequency of insertion of potentially isolated polar sites are important factors in determining the character of hydration. In the extreme concave case of a hydrophobic enzymatic active site, the insertions of polar sites would be expected to be critically important. Whether the surface topography and the addition (or deletion) of polar insertions is indeed used to modulate the character and strength of the hydrophobic interaction in biomolecular systems remains to be determined, but the present results suggest that this new feature may play an important functional role. □

## Methods

Non-bonded Lennard–Jones and coulombic interactions were calculated using atomic pairwise additive potentials. Water was described by the simple point charge (SPC) model<sup>18</sup> and water–protein interactions via optimized potentials in liquid simulations (OPLS) using SPC charges<sup>19</sup>. A cubic periodic boundary condition and a spherical cut-off ( $12.0 \text{ \AA}$ ) minimum image convention (for interactions) were applied. Trajectories were propagated using the Verlet algorithm<sup>20</sup>, and water internal geometry was maintained by using the SHAKE algorithm<sup>21</sup>. Only polar hydrogen atoms were represented and accounted for explicitly. The total number of explicit solute atoms/extended atoms is 510. The surrounding solvent was set up from equilibrated bulk water such that each solute atom is at least  $7 \text{ \AA}$  from the boundary of the central box; a cube edge of  $51.998 \text{ \AA}$  in length containing 4,420 water molecules results. The water molecule closest to each charged centre of the ten charged side chains was replaced with a chloride ion at the beginning of the simulation, yielding a charge-neutral system. Each time step was 2 fs, and configurations were saved at every 10 steps.

Received 28 August 1997; accepted 2 February 1998.

- Teeter, M. M. Water–protein interactions: Theory and experiment. *Annu. Rev. Biophys. Biophys. Chem.* **20**, 577–600 (1991).
- Blokzijl, W. & Engberts, J. B. F. N. Hydrophobic effects. Opinions and facts. *Angew. Chem. Int. Edn Engl.* **32**, 1545–1579 (1993).
- Lee, C. Y., McCammon, J. A. & Rossky, P. J. The structure of liquid water at an extended hydrophobic surface. *J. Chem. Phys.* **80**, 4448–4455 (1984).
- Du, Q., Freysz, E. & Shen, Y. R. Surface vibrational spectroscopic studies of hydrogen bonding and hydrophobicity. *Science* **264**, 826–828 (1994).
- Kuhn, L. A. *et al.* The interdependence of protein surface topography and bound water molecules revealed by surface accessibility and fractal density measures. *J. Mol. Biol.* **228**, 13–22 (1992).
- Sharp, K. A., Nicholls, A., Fine, R. F. & Honig, B. Reconciling the magnitude of the microscopic and macroscopic hydrophobic effects. *Science* **252**, 106–109 (1991).
- Sanchez-Ruiz, J. M. Molecular-size corrections to the strength of the hydrophobic effect: A critical review. *Eur. Biophys. J.* **24**, 261–274 (1996).
- Dempsey, C. E. The actions of melittin on membranes. *Biochim. Biophys. Acta* **1031**, 143–161 (1990).
- Terwilliger, T. C. & Eisenberg, D. The structure of melittin. I. Structure determinations and partial refinement. *J. Biol. Chem.* **257**, 6010–6015 (1982).
- Clarke, A. R. & Lund, P. A. in *The Chaperonins* Ch. 7 (ed. Ellis, R. J.) 167–212 (Academic, San Diego, 1996).



**Figure 3** Comparison of proximal solvent energetics. **a**, Solvent–solvent binding energies ( $E_{\text{bww}}$ ), and **b**, number of hydrogen bonds ( $n_{\text{hb}}$ ) of proximal water molecules around selected residues (these are indicated as in Fig. 2a). Each thin horizontal line is the average of the results obtained for the ten individual time segments. For Leu 13b and Val 8a, separate averages of  $E_{\text{bww}}$  and  $n_{\text{hb}}$  for each individual time segment, are also shown (dashed lines). Bulk values for  $E_{\text{bww}}$  and  $n_{\text{hb}}$  are calculated from water molecules that have a distance of more than  $7 \text{ \AA}$  from any protein atom. These values are computed from the last time segment only.



11. Bernstein, F. C. *et al.* The protein data bank: A computer-based archival file for macromolecular structures. *J. Biol. Chem.* **112**, 535–542 (1977).
12. Kitao, A., Hirata, F. & Go, N. The effects of solvent on the conformation and collective motions of protein: Normal mode analysis and molecular dynamics simulations of melittin in water and in vacuum. *Chem. Phys.* **158**, 447–472 (1991).
13. Mehrotra, P. K. & Beveridge, D. L. Structural analysis of molecular solutions based on quasi-component distribution functions. Application to  $[H_2CO]_{aq}$  at 25°C. *J. Am. Chem. Soc.* **102**, 4287–4294 (1980).
14. Zichi, D. A. & Rossky, P. J. The equilibrium solvation structure for the solvent-separated hydrophobic bond. *J. Chem. Phys.* **83**, 797–808 (1985).
15. Rossky, P. J. & Karplus, M. Solvation. A molecular dynamics study of a dipeptide in water. *J. Am. Chem. Soc.* **101**, 1913–1937 (1979).
16. Pascual-Ahuir, J. L., Silla, E. & Tunon, I. GEPOL: An improved description of molecular surfaces. III. A new algorithm for the computation of the solvent-excluding surface. *J. Comput. Chem.* **15**, 1127–1138 (1994).
17. Freire, E., Murphy, K. P., Sanchez-Ruiz, J. M., Galiseo, M. L. & Privalov, P. L. The molecular basis of cooperativity in protein folding. Thermodynamic dissection of interdomain interactions in phosphoglycerate kinase. *Biochemistry* **31**, 250–256 (1992).
18. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F. & Hermans, J. in *Intermolecular Forces* (ed. Pullman, B.) 331–342 (Reidel, Dordrecht, 1981).
19. Jorgensen, W. L. & Tirado-Rives, J. The OPLS potential functions for proteins. Energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **110**, 1657–1671 (1988).
20. Verlet, L. Computer “experiments” on classical fluids. I. Thermodynamical properties of Lennard–Jones molecules. *Phys. Rev.* **159**, 98–103 (1967).
21. Ryckaert, J.-P., Cicotti, G. & Berendsen, H. J. C. Numerical integration of the Cartesian equations of motion of a system with constraints: Molecular dynamics of *n*-alkanes. *J. Comput. Phys.* **23**, 327–341 (1977).
22. Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281–296 (1991).

**Acknowledgements.** We acknowledge contributions from W. S. Sheu and T. S. Cohen in the early stages of this work. This work was supported by the NIH.

Correspondence and requests for materials should be addressed to P.J.R. (e-mail: rossky@mail.utexas.edu).

## Millennial-scale climate instability during the early Pleistocene epoch

M. E. Raymo\*, K. Ganley\*, S. Carter\*, D. W. Oppo† & J. McManus†

\* Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

† Department of Geology and Geophysics, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

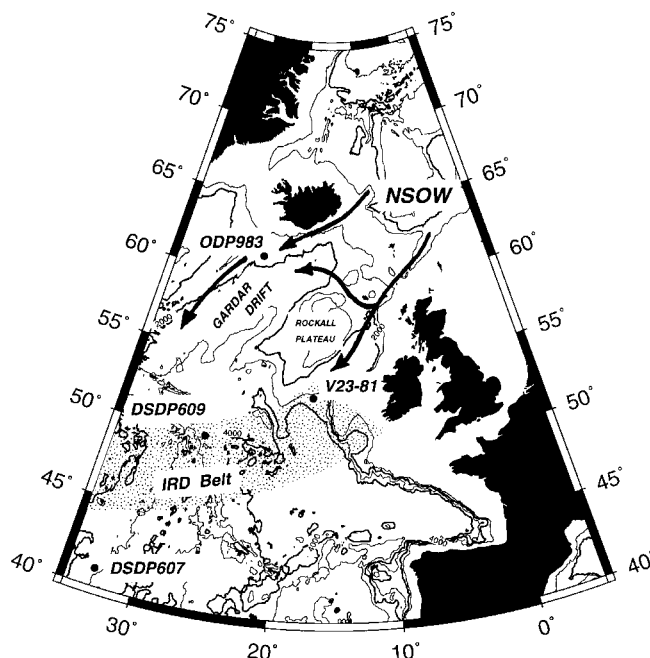
Climate-proxy records of the past 100,000 years show that the Earth's climate has varied significantly and continuously on timescales as short as a few thousand years (refs 1–7). Similar variability has also recently been observed for the interval 340–500 thousand years ago<sup>8</sup>. These dramatic climate shifts, expressed most strongly in the North Atlantic region, may be linked to—and possibly amplified by—alterations in the mode of ocean thermohaline circulation<sup>4–9</sup>. Here we use sediment records of past iceberg discharge and deep-water chemistry to show that such millennial-scale oscillations in climate occurred over one million years ago. This was a time of significantly different climate boundary conditions; not only was the early Pleistocene epoch generally warmer, but global climate variations were governed largely by changes in Earth's orbital obliquity. Our results suggest that such millennial-scale climate instability may be a pervasive and long-term characteristic of Earth's climate, rather than just a feature of the strong glacial–interglacial cycles of the past 800,000 years.

Much has been learned about natural climate variability through the study of ice and sediment cores. In Greenland ice cores, variations of 5–8 °C in mean annual air temperature, called Dansgaard–Oeschger cycles, occur with a rough 1–3 kyr spacing throughout the last glacial cycle<sup>1</sup>. Within deep-sea cores, we observe correlative variations in sea surface temperature (SST), iceberg discharge, and possibly thermohaline circulation<sup>2–9</sup>. The origin of these millennial-scale climate cycles is not well understood; they have been attributed to forcing mechanisms internal to the climate

system, such as ice-sheet dynamics or ocean–atmosphere interactions<sup>4,10,11</sup>, as well as to mechanisms external to the climate system, such as solar variability<sup>12</sup> or combination tones of the main Milankovitch cycles<sup>13,14</sup>. Here we examine the characteristics of climate variability within this frequency band under a markedly different climate regime when glacial episodes were less severe<sup>15</sup>, atmospheric CO<sub>2</sub> concentrations may have been higher<sup>15–17</sup>, and climate over much of the globe (including ice sheets) varied almost exclusively at the 41-kyr period of orbital obliquity<sup>15,18–20</sup>.

To recover such old sediment, we used the unique deep recovery capabilities of the drilling vessel *JOIDES Resolution*. To obtain the required temporal resolution, we drilled the rapidly accumulating Gardar sediment drift located south of Iceland (Fig. 1). Sedimentation rates at this location are enhanced by the horizontal advection of fine (<63-μm) material suspended in a nepheloid layer associated with Norwegian Sea Overflow Water spilling over the Iceland–Faeroe–Scotland ridge<sup>21</sup>. Hence, these sediments record physical and chemical changes in the downwelling limb of the global thermohaline ‘conveyor belt’, changes which strongly influence the circum-Atlantic (and perhaps global) climate<sup>2,4</sup>.

Ocean Drilling Program (ODP) Site 983, recovered at 1,983 m water depth, accumulated continuously throughout the Pleistocene at rates >13 cm kyr<sup>−1</sup> (ref. 22). Oxygen isotope measurements on the benthic foraminifera genus *Cibicides* (primarily *C. wuellerstorfi*) were made every ~38 cm (or ~2,500 yr) for a 27-m section of core spanning the early Pleistocene (Fig. 2). The presence of the Jaramillo (top and bottom) and Cobb Mountain magnetic reversals, diagnostic biostratigraphic datums (including the last occurrence of *Gephyrocapsa* A-B; Fig. 2), and shipboard lithostratigraphic data<sup>22</sup> allowed us to correlate our record to a δ<sup>18</sup>O reference curve from DSDP Site 607<sup>15</sup> (Fig. 2a). We were then able to assign marine isotope stage (MIS) designations<sup>15,20</sup> to the site 983 δ<sup>18</sup>O record (Fig. 2b). Ages for the mid-points of oxygen isotope stages, derived from the astronomically tuned timescale of Shackleton *et al.*<sup>23</sup>, are given



**Figure 1** Location of study area. Map shows the location of core investigated in this study (ODP Site 983 at 60° 24' N, 23° 38' W and 1,983 m water depth) as well as DSDP sites 607 and 609, and V23-81 (ref. 3). Site 983 is located near the head of the Gardar drift which extends to the south-south-west, following the path of deep overflows (Norwegian Sea Overflow Water, NSOW, as shown by arrows) from the Iceland–Faeroe–Scotland ridge<sup>21</sup>. The location of the main ice-rafted debris (IRD) belt as determined for the last glacial cycle is shown by a stippled pattern.

in Fig. 3. Most of the  $\delta^{18}\text{O}$  variance over this interval falls within the range of stage 5a–5d levels at this site<sup>24</sup>.

To study the history of iceberg discharge to this region, the number of lithic fragments ( $>150\text{ }\mu\text{m}$ ) per g of total sediment<sup>3</sup> was determined at orbital-scale resolution for the interval spanning MIS 36 to 45. Lithic grains  $>150\text{ }\mu\text{m}$  in size are delivered to the deep sea by the melting of drifting glacial ice<sup>2,3,9,10</sup> and the correlation of high values of lithic grains per g with positive  $\delta^{18}\text{O}$  (Fig. 2b) reflects the expected relationship between periods of ice-sheet growth and iceberg discharge into the North Atlantic.

To examine suborbital variations in iceberg delivery to the subpolar North Atlantic, lithic counts were made at much higher resolution ( $\sim 5\text{ cm}$  spacing or 380 yr) within MIS 40 and 44 (Fig. 3a and b). Oxygen- and carbon-isotope data were measured for every sample where enough *Cibicidoides* could be found ( $>15\text{ }\mu\text{g}$  of carbonate). The record has noticeably poorer temporal resolution in the early glacial intervals due to a paucity of benthic foraminiferal specimens at these times. During each of the climate cycles studied, we observe episodic increases in the delivery of lithic grains to the site with the final lithic peak within each cycle occurring near the mid-point of the deglacial interval as defined by the  $\delta^{18}\text{O}$  record. Intervals between lithic peaks typically have  $<1$  lithic grain per g of sediment (and many horizons have no lithic grains in the entire sample). Hence, unless one invokes unreasonably large sedimentation rate changes, the lithic data must be recording the episodic input of ice-rafted detritus (IRD) to this site.

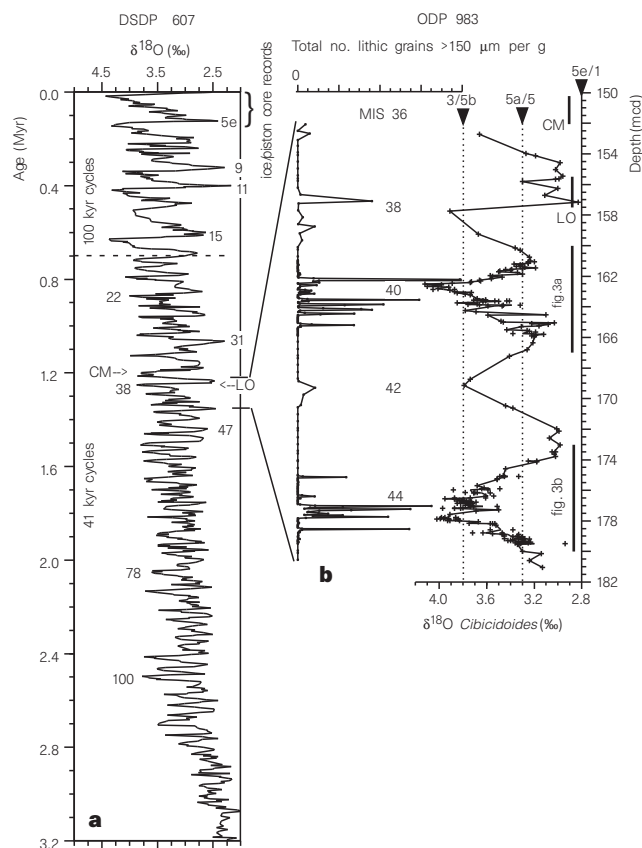
In both MIS 40 and 44, iceberg delivery of lithic material appears to be more pronounced (both in the magnitude and frequency of ice-rafting events) early in the glacial interval, a time also characterized by relatively lower flux of planktonic foraminifera to the sediment (early glacial, EG, interval in Fig. 3). At least six distinct episodes of ice-rafting are observed within glacial stage 40 and five within MIS 44. Given that a complete climate cycle at these times is only  $\sim 41\text{ kyr}$  in duration<sup>20,23</sup>, the ice-rafting episodes, by default, must be occurring more frequently than every 7 kyr (6 episodes  $\times 7\text{ kyr} > 41\text{ kyr}$ ). If one made the typical assumption of constant sedimentation rate between age control points, then the lithic cycles between 163 and 165 m in MIS 40, for instance, would be spaced approximately 2.0–2.5 kyr apart. Assuming a generous range of uncertainty in the sedimentation rates, constrained by the observed degree of symmetry in the glacial–interglacial cycles when plotted to depth (Fig. 2), we estimate these events occur with a periodicity of between 1 and 5 kyr during glacial intervals. Further improvement in chronostratigraphy may come from two sources: progress in understanding secular variations in the strength of the Earth's magnetic field and using these variations as a chronostratigraphic tool<sup>24</sup>; and second, examination of these (and other) specific intervals in different oceanographic settings so as to isolate regional climate versus local sedimentological effects.

Two other proxies for surface water hydrography were examined at this site: percentage *Neogloboquadrina pachyderma* sinistral (left-coiling), a planktonic foraminifera, and the shell  $\delta^{18}\text{O}$  signature of *N. pachyderma* dextral (right-coiling). Both proxies have been shown to covary, during the last glacial cycle, with lithic fragment delivery especially within the main IRD belt<sup>3,7</sup>. However, in the early Pleistocene sections examined here, neither proxy shows a clear millennial-scale signal. This must partly be due to the extreme paucity of planktonic foraminifera within the early glacial interval of maximum lithic delivery (in one case 100% left-coiling *N. pachyderma* reflects the presence of only one planktonic specimen). Also, the location of the main IRD belt relative to Site 983 is unknown for the early Pleistocene, and it has been demonstrated that the surface water SST and meltwater signals are damped away from the main IRD belt (see, for example, ref. 7).

We also note that, in sediments of this age, evolutionary and environmental stability cannot be assumed for the *N. pachyderma* species (see, for example, refs 20, 25). Indeed, MIS 40 falls within an

$\sim 120\text{-kyr}$ -long interval of faunal assemblage variation which has no statistical, or analogous, counterpart in the late Pleistocene ocean<sup>20,25</sup>. This interval is characterized by the near complete absence of *N. pachyderma* sinistral from the North Atlantic, even during peak glacial periods characterized by the same benthic  $\delta^{18}\text{O}$  and IRD signals as intervals (earlier and later) with abundant left-coiling *N. pachyderma*. (Similar anomalous patterns in the distribution of left-coiling *N. pachyderma* are observed before 1.7 Myr; ref. 20). Within the MIS 43–45 interval, *N. pachyderma* sinistral is present at 'normal' levels and reaches a maximum percentage abundance (if a very low absolute abundance) during the early glacial interval of maximum IRD delivery.

From the IRD data we conclude that millennial-scale variations in climate were occurring in the North Atlantic region over one million years ago. In addition, benthic  $\delta^{18}\text{O}$  data, reflecting deep-water chemistry at 1,983 m water depth, present some of the clearest evidence yet found for a link, in the subpolar North Atlantic, between the periodic delivery of icebergs to the region and deep-water circulation changes which can transmit the effects of these regional climate changes around the globe. A number of the lithic peaks observed during MIS 44 and especially during MIS 40 are



**Figure 2** Data from deep-sea sediment cores. **a**, Variation in benthic  $\delta^{18}\text{O}$  at DSDP Site 607 showing pattern of climate variation over past 3.2 Myr (ref. 15). CM indicates position of Cobb Mountain subchron identified in Site 607<sup>20</sup>; LO indicates position of *Gephyrocapsa* A-B datum<sup>21</sup>. Selected marine isotope stages (MIS) are labelled. **b**, 30-m section of early Pleistocene sediment examined from ODP Site 983, including total number of lithics per g of sediment ( $>15\text{ }\mu\text{m}$ ) and  $\delta^{18}\text{O}$  of benthic foraminifera *Cibicidoides*. The depth interval (in metres composite depth; ref. 22) of normal polarity representing the Cobb Mountain (CM) subchron is indicated by shaded bar to left of depth scale. The bar labelled LO indicates the depth range within which the last occurrence of the *Gephyrocapsa* spp. A-B datum occurs. Reference levels for typical  $\delta^{18}\text{O}$  values during late Pleistocene interglacial stages at this site are shown on  $\delta^{18}\text{O}$  curve. Note that 5b/5d levels are also approximately equivalent to high stage 3 levels. The higher density sampling within glacial stages 40 and 44 is shown at an expanded scale in Fig. 3.

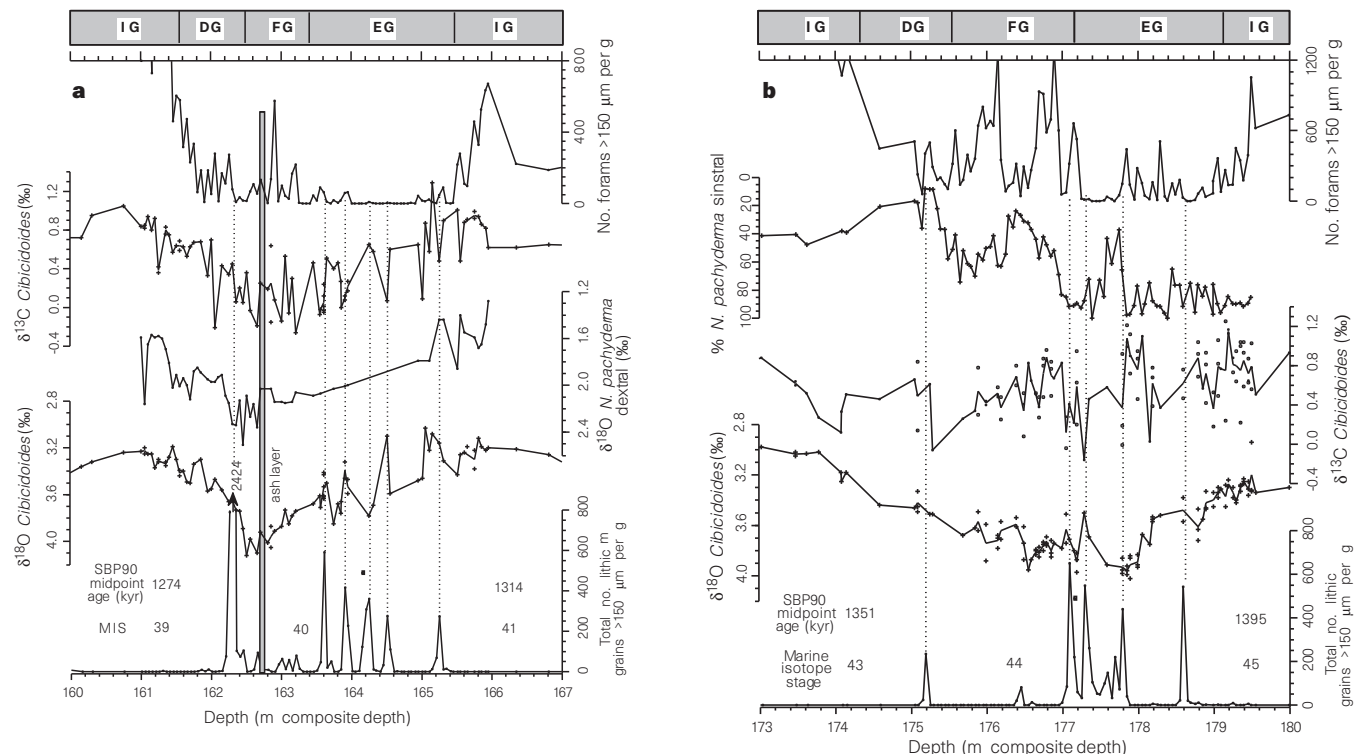
associated with negative excursions in  $\delta^{18}\text{O}$  of up to 0.4‰. As discussed below, these variations in deep-water chemistry do not originate in the surface waters over this site (hence, the fact that a strongly correlative isotopic signal is not observed in the planktonic  $\delta^{18}\text{O}$  record is not unexpected.)

In attempting to explain the origin of these benthic  $\delta^{18}\text{O}$  excursions, we note that a 0.4‰ decrease in  $\delta^{18}\text{O}$  can be caused by a seawater warming of nearly 2 °C or by a decrease in salinity (or some combination of the two changes). If we assume that a salinity decrease is caused by the admixture of melt water and/or runoff to the water mass subducting north of our site (presumably the Norwegian–Greenland Sea), and also that this melt water has an isotopic value of  $-25\text{‰}$ , then subduction of only  $\sim 1.5\%$  glacial melt water by volume would be needed to account for the magnitude of the observed  $\delta^{18}\text{O}$  shifts. This amount of meltwater entrainment in the subducting surface water would cause a corresponding decrease in salinity of 0.5 and in density of  $\sim 0.4$ ). How this less-dense water is then subducted to depth remains uncertain. Whereas Jansen and colleagues<sup>26,27</sup> suggest that brine formation under sea ice in the Norwegian–Greenland Sea is necessary to enhance surface water density to the degree that the meltwater signal can be convected to depth, the modelling results of Lehman *et al.*<sup>28</sup> suggest that the conveyor could entrain and advect a pulse of fresh water directly into the deep ocean.

An alternative hypothesis that has been proposed to explain similar depletions in  $\delta^{18}\text{O}$  observed during the last glacial cycle is the replacement of deep water in this region by warmer Atlantic Intermediate Water<sup>9</sup>. Although speculative, we suggest that our benthic  $\delta^{13}\text{C}$  data is most consistent with the hypothesis that a small amount of highly fractionated melt water is admixed into the subducting surface waters north of Site 983, and that this freshening causes the conveyor to weaken. In the modern ocean, a nearly

constant  $\delta^{13}\text{C}$  composition of  $\sim 1.1\text{‰}$  characterizes water from 1,000 m to greater than 2,500 m depths at the location of Site 983 (ref. 29). During full glacial periods (for example, MIS 2 and 6), deep  $\delta^{13}\text{C}$  values fall to  $\sim 0.5\text{‰}$  at 2,500 m and increase to  $\sim 1.5\text{‰}$  at 1,400 m, with the maximum gradient centred at the depth of Site 983 (ref. 29). The shallower nutrient-depleted water mass is Glacial North Atlantic Intermediate Water (GNAIW) while the deeper nutrient-enriched water mass originates from the Southern Ocean. If a light  $\delta^{18}\text{O}$  excursion was indicative of the presence of warm GNAIW (see, for example, ref. 9) then a positive  $\delta^{13}\text{C}$  shift would be expected. We observed the opposite trend, with  $\delta^{13}\text{C}$  values typically decreasing over the course of an ice-rafting event and remaining low for approximately a millennium afterwards (the clearest expression of this lag is observed at  $\sim 163.6$  and  $163.9$  m depth in the core).

Although such nutrient enrichment is consistent with the subduction of surface water originating in a region of greater than usual iceberg/sea-ice cover, we propose that during an ice-rafting episode, convection in the Norwegian–Greenland Sea begins to weaken owing to increasing melt water at the surface. Entrainment of a small amount of fresh water ( $<2\%$ ) weakens the overflows and limits the depth to which they can sink. This causes the boundary between water from the Southern Ocean and North Atlantic Deep Water to shoal, replacing northern source water with progressively more nutrient-enriched (low  $\delta^{13}\text{C}$ ) southern source water. The unusually high  $\delta^{13}\text{C}$  variability observed during glacial intervals strongly suggests that this site may have been situated near a water-mass boundary in the early Pleistocene as well as during the last glacial. Although this hypothesis is consistent with modelling results<sup>28</sup>, further examination of this site in relation to deeper and shallower sites would be required to determine vertical nutrient distribution and hence the sources and structure of the glacial water masses.



**Figure 3** High-resolution isotopic, faunal and lithologic data from Site 983. **a**, MIS 39 to 41; **b**, MIS 43 to 45. Benthic isotopic measurements were based on *Cibicides*; planktonic isotopic and faunal measurements were on *N. pachyderma*. Ages given for the midpoints of warm stages are after the timescale of Shackleton *et al.* (SBP90)<sup>23</sup>. Vertical dashed lines delineate peaks in the delivery of ice-rafted material to this site. One sample horizon (983A-16-6, 94–98 cm with a lithics per g count of 8,400) was omitted from plot (and interval was labelled as 'ash layer') as it

consisted almost entirely of frothy basaltic glass and probably represents an ash fall (a few non-basaltic lithic grains were observed in this sample, however, and the samples on either side were counted excluding frothy glass). Note also that almost all the benthic isotope analyses for MIS 44 were on single specimens resulting in more replicate analyses and greater variability in the  $\delta^{13}\text{C}$  record, as would be expected from a site located near a strong  $\delta^{13}\text{C}$  gradient. Abbreviations on top axes as follows: IG, interglacial; EG, early glacial; FG, full glacial; DG, deglacial.



We now consider why these rapid climate oscillations occurred in the first place; were they forced (or excited) by processes internal, or external, to the climate system? Here we have shown that millennial-scale variations in iceberg discharge were present over one million years ago and thus occur across a broader spectrum of climate boundary conditions than has been previously recognized. These variations, similar in character and timing to ice-rafting events associated with late Pleistocene Dansgaard-Oeschger cycles, occurred during a climate regime dominated by the obliquity component of orbital forcing and which was too warm to support the growth of large 100-kyr ice sheets<sup>15,20,30</sup>. Correlative variations in the chemical properties of deep North Atlantic water also suggest that glacial melt water associated with periodic ice-sheet mass wasting was entrained in Norwegian Sea Overflow Water, influencing the Atlantic thermohaline 'conveyor belt' and thus, potentially, global climate. □

Received 23 July 1997; accepted 17 February 1998.

- Dansgaard, W. *et al.* in *Climate Processes and Climate Sensitivity* (eds Hansen, J. E. & Takahashi, T.) 288–298 (Am. Geophys. Union, Washington DC, 1984).
- Broecker, W. S. Massive iceberg discharges as triggers for global climate change. *Nature* **372**, 421–424 (1994).
- Bond, G. & Lotti, R. Iceberg discharge into the North Atlantic on millennial time scales during the last glaciation. *Science* **267**, 1005–1010 (1995).
- Broecker, W. S., Bond, G., Klas, M., Bonani, G. & Wolfli, W. A salt oscillator in the glacial Atlantic? I. The concept. *Paleoceanography* **5**, 469–477 (1990).
- Lehman, S. J. & Keigwin, L. D. Sudden changes in North Atlantic circulation during the last deglaciation. *Nature* **356**, 757–762 (1992).
- Behl, R. J. & Kennett, J. P. Brief interstadial events in the Santa Barbara basin, NE Pacific, during the past 60 kyr. *Nature* **379**, 243–246 (1996).
- Oppo, D. W. & Lehman, S. J. Suborbital timescale variability of North Atlantic Deep Water during the past 200,000 years. *Paleoceanography* **10**, 901–910 (1995).
- Oppo, D. W., McManus, J. F. & Cullen, J. C. Abrupt climate change events 500,000 to 340,000 years ago: evidence from subpolar North Atlantic sediments. *Science* **279**, 1335–1338 (1998).
- Rasmussen, T. L., Thomsen, E., Labeyrie, L. & van Weering, T. C. E. Circulation changes in the Faeroe-Scotland Channel correlating with cold events during the last glacial period (58–10 ka). *Geology* **24**, 937–940 (1996).
- MacAyeal, D. R. Binge/purge oscillations of the Laurentide ice sheet as a cause of the North Atlantic Heinrich events. *Paleoceanography* **8**, 775–784 (1993).
- Rahmstorf, S. Rapid climate transitions in a coupled ocean-atmosphere model. *Nature* **372**, 82–85 (1994).
- Suess, H. E. & Linick, T. W. The <sup>14</sup>C record in bristlecone pine of the past 8000 years based on the dendrochronology of the late C. W. Ferguson. *Phil. Trans. R. Soc. Lond. A* **330**, 403–412 (1990).
- Pestiaux, P., van der Mersch, I., Berger, A. & Duplessy, J. C. Paleoclimatic variability at frequencies ranging from 1/10000 to 1/1000 years; evidence for nonlinear behaviour in the climate system. *Clim. Change* **12**, 9–37 (1988).
- Yiou, P. *et al.* High frequency paleovariability in climate and CO<sub>2</sub> levels from the Vostok ice core records. *J. Geophys. Res.* **96**, 20365–20378 (1991).
- Raymo, M. E. in *Start of a Glacial, Proc. Mallorca NATO ARW 207–223* (eds Kukla, G. & Went, E.) (NATO ASI Ser. I, Vol. 3, Springer, Heidelberg, 1992).
- Van Der Burgh, J., Visscher, H., Dilcher, D. L. & Kurschner, W. Paleatmospheric signatures in Neogene fossil leaves. *Science* **260**, 1788–1790 (1993).
- Raymo, M., Grant, B., Horowitz, M. & Rau, G. Mid-Pliocene warmth: stronger greenhouse and stronger conveyor. *Mar. Micropaleontol.* **27**, 313–326 (1996).
- Kukla, G. & An, Z. Loess stratigraphy in central China. *Paleogeogr. Palaeoclimatol. Palaeoecol.* **72**, 203–225 (1989).
- Hooghiemstra, H. Quaternary and upper-Pliocene glaciations and forest development in the tropical Andes: evidence from a long high-resolution pollen record from the sedimentary basin of Bogota, Colombia. *Paleogeogr. Palaeoclimatol. Palaeoecol.* **72**, 11–26 (1989).
- Ruddiman, W. F., Raymo, M. E., Martinson, D. G., Clement, B. M. & Backman, J. Mid-Pleistocene evolution of Northern Hemisphere climate. *Paleoceanography* **4**, 353–412 (1989).
- Wold, C. Cenozoic sediment accumulation on drifts in the northern North Atlantic. *Paleoceanography* **9**, 917–941 (1994).
- Shipboard Scientific Party Proc. ODP Init. Rep. Leg **162**, 139–167 (1996).
- Shackleton, N. J., Berger, A. & Peltier, W. R. An alternative astronomical calibration of the lower Pleistocene timescale based on ODP site 677. *Trans. R. Soc. Edinb. Earth Sci.* **81**, 251–261 (1990).
- Channell, J. E. T., Hodell, D. & Lehman, B. Relative geomagnetic paleointensity and  $\delta^{18}\text{O}$  at ODP Site 983 (Gardar Drift, North Atlantic) since 350 ka. *Earth Planet. Sci. Lett.* (submitted).
- Ruddiman, W. F., Shackleton, N. J. & McIntyre, A. in *North Atlantic Paleoclimatology* (eds Summerhayes, C. P. & Shackleton, N. J.) 155–173 (Spec. Publ. 21, Geol. Soc., London, 1986).
- Veum, T., Jansen, E., Arnold, M., Beyer, I. & Duplessy, J. C. Water mass exchange between the North Atlantic and the Norwegian Sea during the past 28,000 years. *Nature* **356**, 783–785 (1992).
- Jansen, E. & Veum, T. Evidence for two-step deglaciation and its impact on North Atlantic deep-water circulation. *Nature* **343**, 612–616 (1990).
- Lehman, S. J., Wright, D. G. & Stocker, T. F. in *Ice in the Climate System* (ed. Peltier, W. R.) 187–209 (NATO ASI Ser. I, Vol. 12, Springer-Verlag, Berlin, 1993).
- Oppo, D. W. & Lehman, S. J. Mid-depth circulation of the subpolar North Atlantic during the last glacial maximum. *Science* **259**, 1148–1152 (1993).
- Raymo, M. E. The timing of major climate terminations. *Paleoceanography* **12**, 577–585 (1997).
- Wei, W. Calibration of upper Pliocene-lower Pleistocene nannofossil events with oxygen isotope stratigraphy. *Paleoceanography* **8**, 85–100 (1993).

**Acknowledgements.** We thank C. Rosario-Reyes and V. Vieira for assistance, and the Leg 162 Shipboard Scientific Party and crew for their accomplishments and dedication which made the collection of these cores possible. We also thank J. Adkins and T. Dokken for discussions which improved the manuscript. Samples were provided by the ODP with the support of the NSF.

Correspondence and requests for materials should be addressed to M.E.R. (e-mail: raymo@mit.edu).

## Iron partitioning in a pyrolite mantle and the nature of the 410-km seismic discontinuity

Tetsuo Irifune & Maiko Isshiki

Department of Earth Sciences, Ehime University, Matsuyama 790, Japan

Pyrolite<sup>1</sup> is a hypothetical mixture of distinct minerals which is widely believed to represent the composition of the Earth's mantle. The main pressure-induced phase transformations of the olivine component of pyrolite occur at about 13.5 GPa ( $\alpha$  to  $\beta$ ) and 24 GPa ( $\gamma$  to MgSiO<sub>3</sub>-rich perovskite + magnesiowüstite)<sup>2,3</sup>, which are thought to be responsible for the seismic discontinuities at 410 and 660 km depths in the mantle. Recent seismological studies, however, have demonstrated that the 410-km seismic discontinuity is sharper in some areas than that expected from the  $\alpha$  to  $\beta$  transformation in mantle olivine with a fixed composition<sup>4–7</sup>. Moreover, some mineral physics studies suggest that the seismic velocity jump at the 410-km discontinuity is inconsistent with that associated with the  $\alpha$  to  $\beta$  transformation in olivine<sup>8,9</sup>. Here we present a phase equilibria study of a material having pyrolite composition at pressures of 6–16 GPa. We found that the iron content in olivine changes significantly with increasing pressure, as a result of the formation of a relatively iron-rich majorite phase at these pressures. This variation in iron content can overcome, or at least reduce, both of the above difficulties encountered with the pyrolite model of mantle composition, by showing that the component mineral systems cannot be treated as separate.

The pyrolite starting material was a powder of sintered oxide mixture, possessing an Mg number (Mg# = 100 Mg/(Mg + Fe)) of 89.0, which was used in our earlier study<sup>10</sup> at 23–28 GPa. A 4:6 mixture of “pyrolite minus olivine” glass<sup>11</sup> and natural olivine of formula (Mg<sub>0.89</sub>Fe<sub>0.11</sub>)<sub>2</sub>SiO<sub>4</sub> was also used in a few runs to check the equilibrium of the phases. The starting material was sealed in an Au<sub>75</sub>Pd<sub>25</sub> capsule to prevent iron loss from the sample. The capsule was embedded in a platinum tube heater with a boron nitride insulator, and the temperature was monitored at the end of the capsule using a W<sub>97</sub>Re<sub>3</sub>–W<sub>75</sub>Re<sub>25</sub> thermocouple without corrections for the effect of pressure on e.m.f. The temperature uncertainty may be of the order of  $\pm 50^\circ\text{C}$ , mainly due to a thermal gradient within the heater.

High-pressure runs were conducted using a multi-anvil system operated in a 2,000-ton press (ORANGE-2000) at Ehime University. The second-stage cube anvils were truncated by 5 mm edge length on the corners (truncation edge length (TEL) = 5.0), and semi-sintered magnesite was used as a pressure medium. Pressure was carefully calibrated against press load using phase transformations in some reference materials at both room temperature (I–II and III–V in Bi, semiconductor–metal in ZnS and GaAs) and high temperature (coesite–stishovite in SiO<sub>2</sub>; olivine ( $\alpha$ )–modified spinel ( $\beta$ ) in Mg<sub>2</sub>SiO<sub>4</sub>), but may suffer uncertainties of the order of 0.2 GPa. Pressure was applied first, and then temperature was increased to the target value. Heating at the target temperature was conducted for several hours, and then the run was quenched. The recovered sample was examined by an electron microprobe and a micro-focus X-ray powder diffractometer.

Experimental conditions and results of the present runs are summarized in Table 1 and Fig. 1. Runs were conducted at 6–16 GPa at temperatures along a representative geotherm<sup>12</sup>. Electron-microprobe analyses using a broad electron beam demonstrated that there was virtually no iron loss throughout the charge in all of these runs. The  $\alpha$  and  $\beta$  phases of olivine crystallized to give

relatively large grains of 5–20  $\mu\text{m}$ . In contrast, pyroxenes and garnet were generally poorly crystallized (1–5  $\mu\text{m}$ ), which led to relatively large uncertainties in the electron-microprobe analyses of these crystals. Thus the compositions of pyroxenes and garnet have uncertainties of respectively  $\pm 0.5$ –1.3 and  $\pm 0.5$ –2.5 in terms of Mg#, while they are normally within  $\pm 0.3$  for olivine and modified spinel.

Here we define the apparent partition coefficient of Fe and Mg between two phases (a and b) as  $K'_{(a/b)} = \{X_a/(1 - X_a)\} / \{X_b/(1 - X_b)\}$ , where  $X_a$  and  $X_b$  denote mole fractions of iron in phases a and b, respectively. The apparent partition coefficients of the coexisting two phases in pyrolite are calculated from Table 1;  $K'_{(\alpha/\text{gt})} = 0.33$ –0.46,  $K'_{(\alpha/\text{cpx})} = 0.75$ –0.85,  $K'_{(\alpha/\text{opx})} = 1.04$ ,  $K'_{(\beta/\text{gt})} = 0.70$ –0.81,  $K'_{(\beta/\text{cpx})} = 1.37$  and  $K'_{(\alpha/\beta)} = 0.50$ , where cpx indicates clinopyroxene, opx orthopyroxene, and gt garnet. These values are generally in accordance with earlier experimental results obtained for simple systems<sup>13,14</sup> and natural peridotitic samples<sup>15,16</sup>. We also note that there are no notable systematic changes in these values as a function of pressure, especially those at pressures between 10–16 GPa, where the run temperatures do not change significantly (1,320–1,430  $^{\circ}\text{C}$ ).

As shown in Table 1, the chemical compositions of  $\alpha$  and  $\beta$  phases

of olivine observed in the run products using the starting material of 'pyrolite minus olivine' glass + natural olivine coincide with those using the sintered oxide starting material, although the compositions of garnet and clinopyroxene in the former runs were not determined unambiguously owing to their very small crystal sizes. The consistency of the results using two different starting materials suggests that chemical equilibrium was achieved during the present run durations.

Figure 2 shows how the Mg#s of the coexisting phases in pyrolite change with increasing pressure. Clearly all of the phases become more magnesium-rich with increasing pressure, particularly at pressures between 10 and 15 GPa. As shown in Fig. 1, the relatively magnesium-poor majorite garnet phase becomes more abundant at these pressures. However, the equilibrium partitioning relations between these coexisting phases (which vary only weakly with pressure, as noted above) require that this gradual dissolution of magnesium-rich pyroxene into the garnet-majorite solid solution be accompanied by transfer of iron from olivine and its polymorphs into the garnet phase, resulting in net magnesium enrichment of all phases.

It has been demonstrated that the 410-km seismic discontinuity is substantially sharper than that expected from the  $\alpha$  to  $\beta$  phase

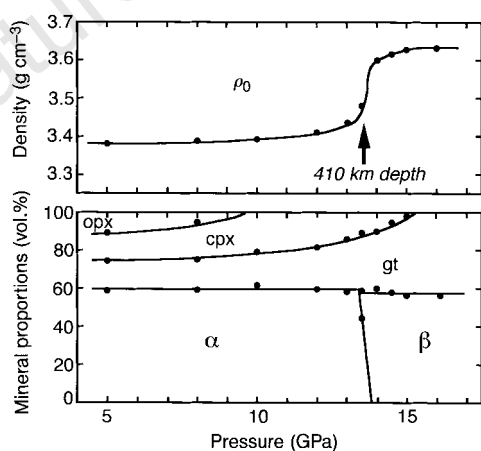
**Table 1** Experimental conditions and results

| Run no.  | Press.<br>(GPa) | Temp.<br>( $^{\circ}\text{C}$ ) | Time<br>(min) | 100Mg/(Mg + Fe) in coexisting phases |         |       |      |      |
|----------|-----------------|---------------------------------|---------------|--------------------------------------|---------|-------|------|------|
|          |                 |                                 |               | $\alpha$                             | $\beta$ | gt    | cpx  | opx  |
| OS-153   | 6               | 1,100                           | 300           | 90.5                                 | —       | 78.1  | 88.8 | 90.8 |
| OS-148   | 8               | 1,250                           | 300           | 91.6                                 | —       | 77.8  | 89.3 | 91.9 |
| OS-138   | 10              | 1,320                           | 300           | 91.4                                 | —       | 81.3  | 88.8 | —    |
| OS-141   | 12              | 1,350                           | 300           | 92.2                                 | —       | 84.5  | 89.9 | —    |
| OS-182a  | 13              | 1,370                           | 180           | 92.3                                 | —       | 83.9  | 90.3 | —    |
| OS-182b* | 13              | 1,370                           | 180           | 92.7                                 | —       | 86.1† | NA   | —    |
| OS-195   | 13.5            | 1,380                           | 180           | 92.3                                 | 85.7    | 83.2  | 91.1 | —    |
| OS-145   | 14              | 1,400                           | 150           | —                                    | 89.1    | 85.1  | 91.8 | —    |
| OS-199   | 14.5            | 1,400                           | 180           | —                                    | 89.7    | 85.1  | NA   | —    |
| OS-189a  | 15              | 1,410                           | 180           | —                                    | 90.8    | 88.8  | NA   | —    |
| OS-189b* | 15              | 1,410                           | 180           | —                                    | 90.4    | 90.2† | NA   | —    |
| OS-165   | 16              | 1,430                           | 180           | —                                    | 91.1    | 88.0  | —    | —    |

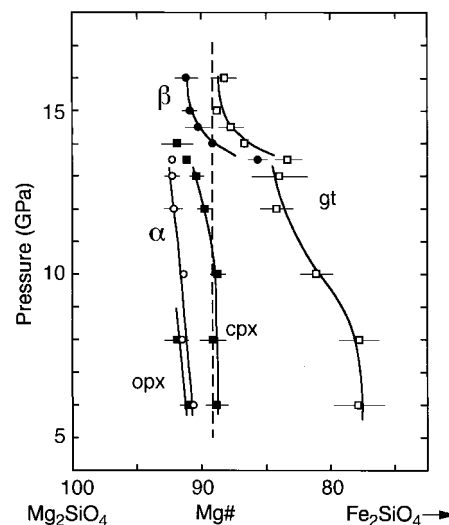
NA, not available due to small grain size and minor quantity, but should be present on the basis of mass-balance considerations and/or X-ray diffraction observations.

\* 'Pyrolite minus olivine' + natural olivine starting material was used.

† Crystals of garnet and clinopyroxene were so small that these values may suffer some uncertainty due to beam overlap in electron microprobe analyses.



**Figure 1** Mineral proportion and zero-pressure density ( $\rho_0$ ) changes in pyrolite as a function of pressure, estimated on the basis of the present experiments. The mineral proportion changes are virtually the same as that estimated by a combination of the experimental results on a 'pyrolite minus olivine' composition and the phase diagram of  $\text{Mg}_2\text{SiO}_4$ - $\text{Fe}_2\text{SiO}_4$  olivine<sup>11</sup>. The zero-pressure densities also coincide with those of the earlier indirect study within the uncertainty ( $\sim \pm 0.02 \text{ g cm}^{-3}$ ) of the calculation (for example,  $\rho_0 = 3.39$  and  $3.64 \text{ g cm}^{-3}$  at 8 and 15 GPa, respectively, in the present study, while they were 3.39 and  $3.65 \text{ g cm}^{-3}$  in the earlier study<sup>11</sup>).  $\alpha$ , olivine;  $\beta$ , modified spinel; opx, orthopyroxene; cpx, clinopyroxene; gt, (majorite) garnet.



**Figure 2** Variations of Mg# (that is,  $100\text{Mg}/(\text{Mg} + \text{Fe})$ ) with pressure in coexisting phases. The dashed line indicates the Mg# value for pyrolite. Filled squares are pyroxenes, open squares are garnet, while circles represent olivine (open) and modified spinel (filled). Horizontal bars represent the analytical uncertainties of the chemical compositions.

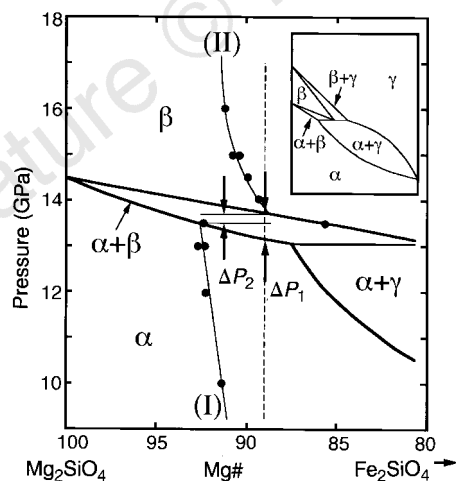
transformation in mantle olivine in some localities<sup>4–7</sup>, although a few recent studies<sup>17,18</sup> indicated that these results may not necessarily be contradictory. Seismological studies<sup>4,6</sup> suggest the transformation interval should be less than 0.1–0.2 GPa, whereas the  $\alpha + \beta$  region for olivine with a fixed  $(\text{Mg}_{0.89}\text{Fe}_{0.11})_2\text{SiO}_4$  composition is estimated<sup>2,3</sup> to be as large as 0.6 GPa. Some mechanisms have been proposed to explain this discrepancy; these include the metastable phase transformation in olivine<sup>19</sup>, the possible existence of melt at the top of the mantle transition region<sup>20</sup>, and the effect of trace amounts of water on the transformation interval<sup>21</sup>. These studies assume that the compositions of mantle olivine and its high-pressure form do not change on the phase transformation. However, our study clearly demonstrates that they do change with pressure as a result of the chemical interaction between the coexisting phases, particularly due to the formation of majorite garnet at pressures between 10 and 15 GPa. This may reduce the transformation interval to  $\sim 0.2$  GPa as seen in Fig. 3, which is consistent with that deduced from seismic observations in certain areas. A small local change in olivine composition in the actual mantle may cause some variation in this interval, which should produce variations in the sharpness of the 410-km discontinuity at different localities<sup>4–7</sup>. Thus the seismologically observed sharpness of the 410-km discontinuity may be reconciled with that expected from the  $\alpha$ – $\beta$  transformation in the pyrolite mantle without invoking any additional mechanisms such as proposed earlier<sup>19–21</sup>.

It has also been common practice in mineral physics to assume a fixed composition for mantle olivine in order to calculate density and seismic-velocity profiles as a function of depth. The density and velocity profiles thus calculated for pyrolite or peridotite compositions are generally in agreement with the seismologically derived profiles throughout the upper mantle and the mantle transition region<sup>11,22–24</sup>, although compositions with less olivine content also match the observations<sup>25</sup>. One notable discrepancy between the observed and calculated profiles is the magnitude of the seismic velocity jump at 410 km depth, as pointed out by some mineral-physics studies<sup>8,9,25</sup>. These studies suggest that olivine content in the

upper mantle should be in the range  $\sim 30$ – $50$  vol.% to match the observed velocity jump at 410 km depth, although recent experimental studies concluded that pyrolite or peridotite compositions with  $\sim 60$  vol.% olivine are still acceptable<sup>22,26,27</sup>.

We have shown that olivine becomes Mg-rich with increasing pressure, possessing Mg#s of 92–93 at 13.5 GPa, equivalent to 410 km depth, while modified spinel has a composition of Mg# = 88–89 immediately after the transformation is completed (Fig. 3). Recent measurements of elastic moduli of olivine at high pressure showed that both bulk modulus and shear modulus of pure forsterite are similar to, or slightly higher than, those of iron-bearing (Mg# = 90) olivine, particularly at pressures above 10 GPa (ref. 28). The density of olivine becomes substantially higher with increasing iron content, and thus the seismic velocities of iron-poor olivine should be notably higher than those of iron-rich olivine. The longitudinal wave velocity ( $V_p$ ) and transverse wave velocity ( $V_s$ ) jumps associated with the  $\alpha$  (Mg# = 92.5) to  $\beta$  (Mg# = 88.5) transformation in the pyrolite composition calculated from a mineral-physics database<sup>25</sup> at ambient conditions ( $\Delta V_p = 11.6\%$  and  $\Delta V_s = 12.1\%$ ) are smaller by  $\sim 10\%$  than those expected in the isochemical transformation of olivine with Mg# =  $\sim 89$  ( $\Delta V_p = 12.8\%$  and  $\Delta V_s = 13.8\%$ ). This means that the proportion of olivine in the mantle compositions should be increased by  $\sim 4$ – $5$  vol.% in the previous estimations based on the isochemical phase transformation of olivine. Moreover, recent measurements of  $V_p$  and  $V_s$  at high pressure for the  $\alpha$  and  $\beta$  phases of olivine showed that the velocity jumps decrease with increasing pressure<sup>26</sup>, which requires a larger proportion of olivine than those estimated on the basis of the velocity jump at ambient conditions.  $\square$

Received 26 August 1997; accepted 17 February 1998.



**Figure 3** The observed variations of Mg# in olivine (I) and modified spinel (II) in pyrolite, plotted on the phase diagram of  $\text{Mg}_2\text{SiO}_4$ – $\text{Fe}_2\text{SiO}_4$  at 1,400°C. The  $\alpha$ – $\beta$  phase boundaries were drawn so as to match the chemical compositions of the coexisting  $\alpha$  and  $\beta$  phases observed at 13.5 GPa and also by fixing the transformation pressure for the  $\text{Mg}_2\text{SiO}_4$  composition to 14.5 GPa (ref. 29). These boundaries are generally consistent with the results of experimental and thermodynamic studies for the system  $\text{Mg}_2\text{SiO}_4$ – $\text{Fe}_2\text{SiO}_4$  (refs 2, 3), but they should be regarded as tentative because we do not know the effects of the other minor components on the phase boundaries. It is seen that the pressure interval of the  $\alpha + \beta$  region for pyrolite composition ( $\Delta P_2$ ) is about one-third of that expected in olivine with a fixed  $(\text{Mg}_{0.89}\text{Fe}_{0.11})_2\text{SiO}_4$  composition ( $\Delta P_1 \approx 0.6$  GPa).

- Ringwood, A. E. In *Advances in Earth Sciences* (ed. Hurley, P. M.) 287–356 (MIT Press, Cambridge, MA, 1966).
- Katsura, T. & Ito, E. The system  $\text{Mg}_2\text{SiO}_4$ – $\text{Fe}_2\text{SiO}_4$  at high-pressures and temperatures: precise determination of stabilities of olivine, modified spinel, and spinel. *J. Geophys. Res.* **94**, 15663–15670 (1989).
- Akaogi, M., Ito, E. & Navrotsky, A. J. Olivine-modified spinel-spinel transitions in the system  $\text{Mg}_2\text{SiO}_4$ – $\text{Fe}_2\text{SiO}_4$ : Calorimetric measurements, thermochemical calculation, and geophysical application. *J. Geophys. Res.* **94**, 15671–15685 (1989).
- Benz, H. M. & Vidale, J. E. Sharpness of upper mantle discontinuities determined from high-frequency reflections. *Nature* **365**, 147–150 (1993).
- Yamazaki, A. & Hirahara, K. The thickness of upper mantle discontinuities, as inferred from short-period J-array data. *Geophys. Res. Lett.* **22**, 2557–2560 (1995).
- Vidale, J. E., Ding, X. Y. & Grand, S. P. The 410-km-depth discontinuity: a sharpness estimate from near-critical reflections. *Geophys. Res. Lett.* **22**, 2557–2560 (1995).
- Neele, F. Sharp 400-km discontinuity from short period P reflections. *Geophys. Res. Lett.* **23**, 419–422 (1996).
- Duffy, T. S. Elasticity of forsterite to 16GPa and the composition of the upper mantle. *Nature* **378**, 170–173 (1995).
- Zha, C. S. et al. Single-crystal elasticity of  $\beta$ - $\text{Mg}_2\text{SiO}_4$  to the pressure of the 410-km seismic discontinuity in the Earth's mantle. *Earth Planet. Sci. Lett.* **147**, 9–15 (1997).
- Irfune, T. Absence of an aluminous phase in the upper part of the Earth's lower mantle. *Nature* **370**, 131–133 (1994).
- Irfune, T. An experimental investigation of the pyroxene-garnet transformation in a pyrolite composition and its bearing on the constitution of the mantle. *Phys. Earth Planet. Inter.* **45**, 324–336 (1987).
- Ito, E. & Katsura, T. A temperature profile of the mantle transition zone. *Geophys. Res. Lett.* **16**, 425–428 (1989).
- Nishizawa, O. & Akimoto, S. Partition of magnesium and iron between olivine and spinel and between pyroxene and spinel. *Contrib. Mineral. Petrol.* **41**, 217–230 (1973).
- Akimoto, S. et al. In *The Physics and Chemistry of Minerals and Rocks* (ed. Strens, R. G. J.) 327–363 (Wiley, London, 1976).
- Akaogi, M. & Akimoto, S. High-pressure phase equilibria in a garnet lherzolite, with special reference to  $\text{Mg}^{2+}$ – $\text{Fe}^{2+}$  partitioning among constituent minerals. *Phys. Earth Planet. Inter.* **19**, 31–51 (1979).
- Takahashi, E. & Ito, E. In *High-Pressure Research in Mineral Physics* (eds Manghni, M. H. & Syono, Y.) 427–437 (Terra Publ./Am. Geophys. Union, Tokyo, 1987).
- Helffrich, G. & Bina, C. R. Frequency dependence of the visibility and depths of mantle seismic discontinuities. *Geophys. Res. Lett.* **21**, 2613–2616 (1994).
- Stixrude, L. Structure and sharpness of phase transitions and mantle discontinuities. *J. Geophys. Res.* **102**, 14835–14852 (1997).
- Solomatov, V. S. & Stevenson, D. J. Can sharp seismic discontinuities be caused by non-equilibrium phase transformations? *Earth Planet. Sci. Lett.* **125**, 267–279 (1994).
- Revernaugh, J. & Sipkin, S. A. Seismic evidence for silicate melt at the 410 km mantle discontinuity. *Nature* **369**, 474–476 (1994).
- Wood, B. J. The effect of  $\text{H}_2\text{O}$  on the 410-kilometer seismic discontinuity. *Science* **268**, 74–76 (1995).
- Weidner, D. J. A mineral physics test of a pyrolite mantle. *Geophys. Res. Lett.* **12**, 417–420 (1985).
- Bina, C. R. & Wood, B. J. The 400-km seismic discontinuity and the proportion of olivine in the Earth's upper mantle. *Nature* **324**, 449–451 (1986).
- Gwanmesia, G. D., Rigden, S., Jackson, I. & Liebermann, R. C. Pressure dependence of elastic wave velocity for  $\beta$ - $\text{Mg}_2\text{SiO}_4$  and the composition of the Earth's mantle. *Science* **250**, 794–797 (1990).
- Duffy, T. S. & Anderson, D. L. Seismic velocities in mantle minerals and the mineralogy of the Earth as determined from seismic profiles. *J. Geophys. Res.* **94**, 1895–1912 (1989).



26. Li, B., Gwanmesia, G. D. & Liebermann, R. C. Sound velocities of olivine and beta polymorphs of  $\text{Mg}_2\text{SiO}_4$  at Earth's transition zone pressure. *Geophys. Res. Lett.* **23**, 2259–2262 (1996).
27. Rigden, S. M. *et al.* Spinel elasticity and the seismic structure of the transition zone of the mantle. *Nature* **354**, 143–145 (1991).
28. Zha, C. S., Duffy, T. S., Mao, H. K. & Hemley, R. J. in *High-Pressure Temperature Research: Properties of Earth and Planetary Materials* (eds. Manghnani, M. H. & Yagi, T., in the press).
29. Morishima, T. *et al.* The phase boundary between  $\alpha$ - and  $\beta$ - $\text{Mg}_2\text{SiO}_4$  determined by in situ X-ray observation. *Science* **265**, 1202–1203 (1994).

**Acknowledgements.** We thank T. Inoue for discussions, K. Fukuhama for initial experiments, and C. R. Bina for comments. This work was supported by Grant-in-Aids for Scientific Research from Monbusho and JSPS.

Correspondence and requests for materials should be addressed to T.I. (e-mail: irifune@dpc.ehime-u.ac.jp).

## Oldest known sea turtle

Ren Hirayama

Teikyo Heisei University, Uruido 2289, Ichihara, Chiba 290-01, Japan

Reptiles constitute a primarily terrestrial assemblage, but several groups returned to the marine environment after the first appearance of reptiles in the late Palaeozoic era<sup>1</sup>. Successful diversification of the chelonoid sea turtles, particularly during the Cretaceous period, was perhaps one of the most important events in the history of turtles (and marine reptiles)<sup>2–4</sup>. The fossil record of chelonoids before the Late Cretaceous has been poorly documented. Here I report the discovery of an exceptionally well-preserved skeleton of the oldest known chelonoid, from the Early Cretaceous stage (about 110 million years before the present)<sup>5–7</sup> of eastern Brazil. This specimen represents a new taxon, extending the history of chelonoids by 10 million years,

and it sheds new light on the early evolution of the group. The limb of the specimen is a relatively primitive paddle, which still possesses movable digits as in freshwater turtles. However, the skull is specialized in the manner of later chelonoids, with large interorbital foramina that are indicative of huge lachrymal salt glands surrounding the eyes<sup>8,9</sup>. This discovery supports the idea that the establishment of the salt-excreting system, and the occupation of a marine habitat, may have preceded the formation of rigid paddles in the history of chelonoids.

Testudines Linnaeus 1766

Cryptodira Gray 1825

Eucryptodira Gaffney 1975

Chelonioidea Agassiz 1857

Protostegidae Cope 1873

*Santanachelys gaffneyi*, gen. et sp. nov.

**Etymology.** Santana, for the locality; chelys (Greek): turtle; gaffneyi, in honour of Eugene S. Gaffney, an authority on fossil turtles.

**Holotype.** THUg1386 (Teikyo Heisei University, Ichihara, Chiba, Japan), a nearly complete skeleton (Fig. 1).

**Locality.** Near Santana do Cariri, Ceara State, Brazil.

**Horizon.** Romualdo Member, Santana Formation (Early Cretaceous: Late Aptian or Early Albian)<sup>5–7</sup>.

**Diagnosis.** Scute sulci on the dermal roofing elements of skull and the carapace are present; the nasal is much smaller than the prefrontal; the prefrontal–postorbital contact at the orbital rims is absent; the foramen palatinum posterius opens laterally; the lingual ridge of the maxilla is low, and not exposed from lateral ridge; the first thoracic rib is elongated, reaching to the distal end of the first costal; the lateral process of the humerus lacks a medial concavity; the articulations of the first and second metacarpals and



**Figure 1** Holotype of *Santanachelys gaffneyi*, gen. et sp. nov. (THUg1386). The median length of the preserved carapace is 145 mm. **a**, Dorsal view. **b**, Ventral view. **c**, Left lateral view.

digits are movable; the xiphiplastra are as long as they are broad, and are medially attached along their entire length (Figs 1–3).

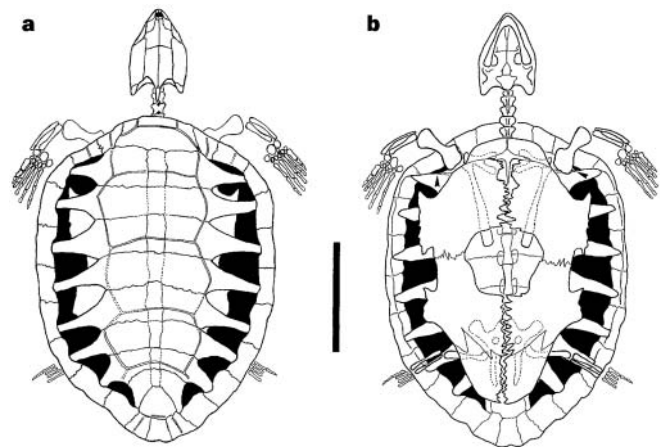
*Santanachelys* differs from all other chelonoids in its possession of an unusually elongate first thoracic rib. It differs from all other protostegids in its possession of large xiphiplastra and movable articulations of first and second metacarpals and digits, and in lacking a medial concavity of the lateral process of the humerus<sup>2,4</sup>. It differs from *Rhinochelys*, a protostegid known primarily from isolated skulls, in possessing much smaller nasal bones<sup>2,10</sup> and also in lacking an antero-dorsal bulge (Fig. 3a) (refs 2, 10).

The discovery of a chelonoid in the Santana Formation of Brazil extends the range of this group from the Late Albian (where species such as *Rhinochelys* and *Notochelone* have been found) back to the Early Albian or Late Aptian<sup>2,4</sup>. The relationships of *Santanachelys* to other chelonoids and eucryptodires (Fig. 4) show that this is the most primitive member of the protostegid family of Cretaceous sea turtles<sup>2,4,10,11</sup>. This result also implies that chelonoids had diversified into the three families (Protostegidae, Dermochelyidae and Cheloniidae) by the Early Albian or Late Aptian, and that the origin of chelonoids might be traced back much further, possibly to the Neocomian<sup>12</sup>. *Santanachelys* is also the first known protostegid from South America. Despite its small body size (200-mm long as preserved), the highly ossified skeleton, particularly in the carapace and manus, indicates that this individual may have reached at least sub-adult size.

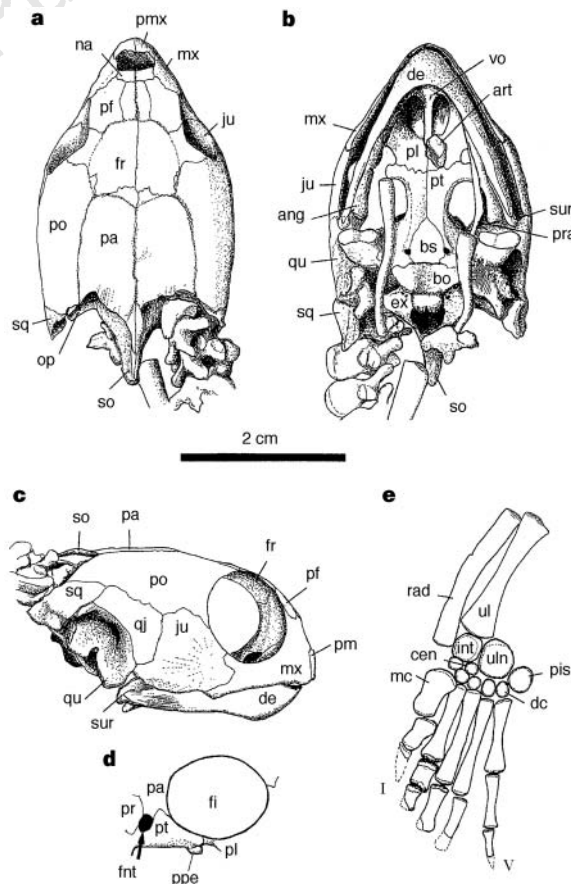
One of the major questions about the chelonoid adaptation to marine environments has concerned the timing of the modification of its limbs into paddle-like, propulsive organs, involving the immobilization of the manus and pes<sup>11,13,14</sup>. The more primitive stage of paddle morphology was hitherto known only in some primitive chelonoids such as *Toxochelys*<sup>11</sup> (in the Late Cretaceous and *Tasbacka*<sup>15</sup> (in the Paleocene), in which movable first and second metacarpals and (short) digits were retained. The discovery of this primitive grade of paddle in *Santanachelys* (Fig. 3e) indicates that the complete formation of paddles, consisting of immobile metacarpals and elongate digits, may have been independently acquired in each of the three families of chelonoids.

Studies of the living chelonoids have revealed another important adaptation to marine life among sea turtles<sup>8,9</sup>. Most marine vertebrates must continuously confront the problem of water loss and salt gain; however, the kidneys of reptiles are unable to handle a large salt influx. Marine reptiles are unique in the variety of cephalic organs that can be used as salt glands. Salt excretion in sea turtles mainly occurs in salt-gland fluids secreted by highly modified lachrymal glands, which are much larger than the brain. The chelonoids are unique among the casichelydian turtles (turtles

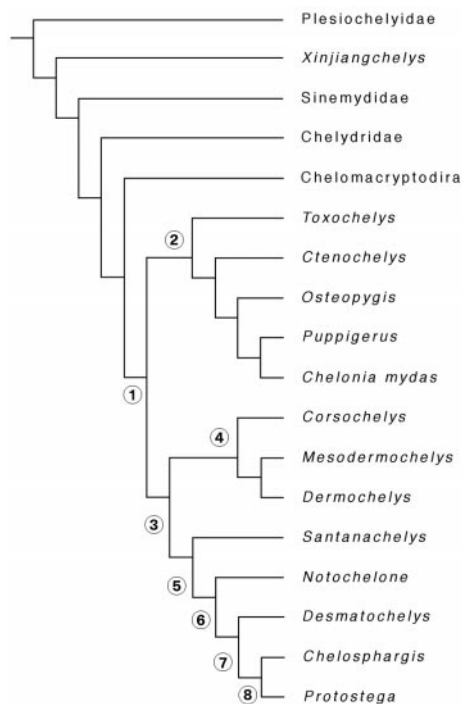
descended exclusively from *Proganochelys*, the most primitive turtle from the Late Triassic) in that they possess a large foramen interorbitale formed by a relatively narrow processus inferior parietalis<sup>16–18</sup>. The processus is completely lost in living *Dermochelys coriacea*<sup>17</sup>, which has particularly large lachrymal glands<sup>19</sup>. The enlargement of the posterior portion of the foramen interorbitale seems to be related to the development of the large lachrymal gland, which lies postero-medial to the eye of sea turtles<sup>19,20</sup>. This portion of the foramen interorbitale is limited by a much broader processus inferior parietalis in most turtles<sup>16,17</sup>. The Jurassic plesiochelyids, primitive eucryptodires once considered to be ancestral chelonoids<sup>16</sup> but later regarded as more basal animals<sup>14,21</sup> (Fig. 4), have a processus that is intermediate in size between those of the chelonoids and those of the other casichelydians<sup>16</sup>. A large foramen interorbitale with a narrow processus inferior parietalis like that in *Santanachelys* (Fig. 3d) is also seen in primitive chelonoids such as *Toxochelys*<sup>2</sup> (Cheloniidae) and *Corsochelys*<sup>22</sup> (Dermochelyidae). This suggests that the large, salt-excreting lachrymal glands were already developed in this primitive sea turtle, and that the evolutionary appearance of the lachrymal gland preceded the formation of rigid paddles in the history of chelonoids. The function of the gland in excreting salt indicates that *Santanachelys* may have lived in a marine habitat (as



**Figure 2** Skeletal reconstruction of *Santanachelys gaffneyi*, (THUG1386). **a**, Dorsal view. **b**, Ventral view, with appendicular skeleton overlying plastron shown by dotted lines. Arrowheads in **b** indicate the distal end of first thoracic ribs. Scale bar, 5 cm.



**Figure 3** *Santanachelys gaffneyi* (THUG1386). The skull in **(a)** dorsal, **(b)** ventral and **(c)** right lateral views is shown, as is **(d)** the right ethmoid region of the skull. **e**, Left forelimb, dorsal view. Abbreviations: ang, angular; art, articular; bo, basioccipital; bs, basisphenoid; cen, centrale; dc, distal-carpal; de, dentary; ex, exoccipital; fi, foramen interorbitale; fr, frontal; fnt, foramen nervi trigemini; int, intermedium; ju, jugal; mc, metacarpal; mx, maxilla; na, nasal; op, opisthotic; pa, parietal; pf, prefrontal; pis, pisiform; pl, palatine; pmx, premaxilla; po, postorbital; ppe, processus pterygoideus externus; pr, pro-otic; pra, prearticular; pt, pterygoid; qj, quadrato-jugal; qu, quadrate; rad, radius; so, supraoccipital; sq, squamosal; sur, surangular; ul, ulna; uln, ulnare; vo, vomer.



**Figure 4** Cladogram showing relationships among selected eucryptodiran turtles. Cheloniioidea is the ingroup; other taxa are outgroups. The analysis is based on 104 characters; the tree length is 208; the consistency index is 0.567; and the retention index is 0.720. Circle numbers indicate nodes or branching points of chelonioids (see Methods).

also indicated by the depositional environment of the fossil)<sup>5,6</sup>, and that the shift in environment preceded the evolution of paddles. □

## Methods

Cladistic analysis was based on the implementation of the software package PAUP (phylogenetic analysis using parsimony) version 3.0, developed by D. L. Swofford. Terminal taxa included in the analysis are the Plesiochelyiidae, *Xinjiangchelys*, Sinemydidae, Chelydridae (including *Platysternon*), Chelomacryptodira (Testudinoidea and Trionychoidea), and Cheloniioidea, including Cheloniidae (*Toxochelys*, *Ctenochelys*, *Osteopygis*, *Puppigerus* and *Chelonia mydas*), Dermochelyiidae (*Corsochelys*, *Mesodermochelys* and *Dermochelys*), and Protostegidae (*Santanachelys*, *Notochelone*, *Desmatochelys*, *Chelosphargis* and *Protostega*). For the list of the 104 osteological characters and the data matrix for the analysis, see Supplementary information. A single tree was obtained by a branch-and-bound search option (Fig. 4).

All characters were coded as reversible, and multistate characters were coded as unordered to avoid *a priori* assumptions of transformation vectors. When one or more taxa are coded as having multiple states, I interpret multiple states as uncertainty rather than polymorphism. Synapomorphies for the numbered nodes of chelonioids (Fig. 4) follow DELTRAN character optimization, because it is slightly more conservative in terms of assigning synapomorphies to clades in a data matrix with a substantial amount of missing data. Synapomorphies also shared unequivocally by ACCTRAN character optimization are italicized. For the complete list of synapomorphies of both DELTRAN and ACCTRAN character optimizations, see Supplementary information.

1. Cheloniioidea. *Processus inferior parietalis* as narrow as the foramen nervi trigemini or absent; foramen stapediale temporale concealed from dorsal by dermal roofing elements; *foramen premaxillae* absent; *paired foramina anterius canalis carotici interni* lie close together; *dorsum sellae* high and separated from *sella turcia* and *anterius canalis carotici interni*; symphyseal ridge of dentary present; first thoracic ribs limited within nuchal width; coracoid as long as or much longer than humerus; humerus longer than femur; lateral process of the humerus located distal to caput humeri; *flat carpal and tarsal elements*; *Third to fifth digits without movable articulations*; *distal portion of costals reduced*; scute

sulci on plastron rudimentary or absent; *entoplastron not tightly sutured with hyoplastron*.

2. Cheloniidae. Orbital direction dorsolateral; *parietal–squamosal contact* present; *upper triturating surface* involving palatines; *apertura narium interna* entirely formed by vomer and palatines, excluding maxillae; ventral surface of basisphenoid with V-shaped crest emerginated from posterior; ventral thin keel at posterior cervical centra present; *ulna–radius contact* through prominent distal rugosities; cervical scute overlying more than half width of nuchal; *neural number nine* by splitting of eighth or seventh neurals.

3. Dermochelyiidae (Dermochelyiidae and Protostegidae). *Medial process of jugal* absent; jugal–pterygoid contact absent; *proportion of cervical centra* as high as wide; eighth cervical centrum longer than seventh; first thoracic vertebra with its anterior articulation facing anteriorly; scapular angle formed by scapular prong and acromion about 110° or more; thyroid fenestra small, subdivided; lateral process of ischium small; shoulder of caput humeri absent; femoral trochanters connected by bony ridge; neural bones rectangular; *plastral index* larger than 100; *entoplastron T-shaped*, with distinct lateral wings.

4. Dermochelyiidae. Cranial scute sulci on dermal roofing elements absent; *processus trochlearis oticum* formed by prootic reduced; ossification of rostrum basisphenoidale reduced; foramen caroticum laterale larger than foramen anterius canalis carotici interni; scute sulci on carapace absent or rudimentary; *plastral fontanelles* between hyo–hypoplastra as large as or much larger than hyoplastron, or hypoplastron; medial contact of xiphiplastra absent.

5. Protostegidae. *Nasal* present; jugal–quadrate contact, excluding quadrado–jugal from lower cheek margin; lower cheek emargination absent; vomer–palatine contact lost, palatines medially meeting; foramen palatinum posterius open postero–laterally; pterygoid extending onto mandibular articular surface of quadrate; foramen posterius canalis carotici interni between pterygoid and basisphenoid; *rodlike rostrum basisphenoidale*; third (or second) cervical centrum biconvex; lateral process of pubis extending anteriorly beyond midline portion of pubis; lateral process of humerus enlarged within the anterior portion of shaft, not easily visible from ventral view; radius with middle portion bent toward anterior; cervical scute overlying more than half the width of nuchal; *thick neurals* with median keel.

6. Unnamed taxon (protostegids excluding *Santanachelys*). Upper triturating surface involving vomer; *high lingual ridge of maxilla* exposed from lateral; lateral process of humerus with a medial concavity; first and second digits modified into paddle without movable articulations; *xiphiplastra* much broader than long, curved medially, with large midline fontanelle.

7. Unnamed taxon. Foramen palatinum posterius absent; pterygoid narrow, C-shaped; *large processus trochlearis oticum* involving deep notch of quadrate; first suprapygial absent.

8. Unnamed taxon. Prefrontal–postorbital contact present; scar for *M. latissimus dorsi* and *teres major* located at middle of shaft; *star-shaped*, hyo–hypoplastra.

Received 17 September 1997; accepted 20 January 1998.

- Carroll, R. L. *Vertebrate Paleontology and Evolution* (Freeman, New York, 1988).
- Hirayama, R. Phylogenetic systematics of chelonoid sea turtles. *Island Arc* 3, 270–284 (1995).
- Hirayama, R. & Chitoku, T. Family Dermochelyiidae (Superfamily Cheloniioidea) from the Upper Cretaceous of North Japan. *Trans. Proc. Palaeontol. Soc. Japan* 184, 597–622 (1996).
- Hirayama, R. in *Ancient Marine Reptiles* (eds Callaway, J. L. & Nicholls, E. L.) 241–257 (Academic, San Diego, 1997).
- Maisey, J. G. in *Santana Fossils: An Illustrated Atlas* (ed. Maisey, J. G.) 11–21 (TFH Publications, Neptune, NJ, 1991).
- Martill, D. M. Fossils of the Santana and Crato Formations, Brazil. *Palaeontol. Assoc. Field Guide Foss.* 5, 1–159 (1993).
- Meylan, P. A. Skeletal morphology and relationships of the Early Cretaceous side-necked turtle, *Araripemys barreroi* (Testudines: Pelomedusoides: Araripemydidae), from the Santana Formation of Brazil. *J. Vert. Paleontol.* 16, 20–33 (1996).
- Schmidt-Nielsen, K. & Fange, R. Salt glands in marine reptiles. *Nature* 183, 783–785 (1958).
- Lutz, P. L. in *The Biology of Sea Turtles* (eds Lutz, P. L. & Musick, J. A.) 343–361 (CRC, New York, 1997).
- Collins, J. I. The chelonian *Rhinochelys* Seeley from the Upper Cretaceous of England and France. *Palaeontology* 13, 355–378 (1970).
- Zangerl, R. The vertebrate fauna of the Selma Formation of Alabama. Part 3. The turtles of the family Protostegidae. Part 4. The turtles of the family Toxochelyiidae. *Fieldiana Geol. Mem.* 3, 61–277 (1953).
- Shaffer, H. B., Meylan, P. & McKnight, M. L. Tests of turtle phylogeny: molecular, morphological, and paleontological approaches. *Syst. Biol.* 46, 235–268 (1997).
- Walker, W. E. Jr in *Biology of the Reptilia* 4 (eds Gans, C. & Parsons, T. S.) 1–100 (Academic, New York, 1973).
- Gaffney, E. S. & Meylan, P. A. in *The Phylogeny and Classification of Tetrapods. 1. Amphibians, Reptiles, Birds* (ed. Benton, M. J.) 157–219 (Clarendon, Oxford, 1988).
- Nessov, L. A. The Paleogene sea turtles of Southern Kazakhstan and the phylogenetic relationships between Toxochelyiidae and the Cheloniidae. *Palaeontol. J.* 21, 73–84 (1987).



16. Gaffney, E. S. Cranial morphology of the European Jurassic turtles *Portlandemys* and *Plesiochelys*. *Bull. Am. Mus. Nat. Hist.* **157**, 489–543 (1976).
17. Gaffney, E. S. Comparative cranial morphology of recent and fossil turtles. *Bull. Am. Mus. Nat. Hist.* **164**, 65–375 (1979).
18. Gaffney, E. S. The comparative osteology of the Triassic turtle *Proganochelys*. *Bull. Am. Mus. Nat. Hist.* **194**, 1–263 (1990).
19. Schumacher, G.-H. Die Kopf- und Halsregion der Leder-schildkröte *Dermochelys coriacea* (Linnaeus 1776). *Abhandl. Akad. Wissensch.* **1972**, 1–60 (1973).
20. Schumacher, G.-H. in *Biology of the Reptilia*. 4 (eds Gans, C. & Parsons, T. S.) 101–199 (Academic, New York, 1973).
21. Rieppel, O. The skull of the Upper Jurassic cryptodire turtle *Thalassermys*, with a reconsideration of the chelonian braincase. *Palaeontogr. A* **171**, 105–140 (1980).
22. Zangerl, R. The vertebrate fauna of the Selma Formation of Alabama. Part 5. An advanced chelonoid sea turtle. *Fieldiana Geol. Mem.* **3**, 61–277 (1960).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from Mary Sheehan at the editorial office of Nature.

**Acknowledgements.** I thank F. Bacchia for field collection of the Santana turtles; E. S. Gaffney, P. E. Meylan and T. Hirayama for comments on manuscript; E. Hooks III for advice on protostegid morphology; D. B. Brinkman for comments on primitive eucryptodires; and N. Kohno for advice on the methodology of phylogenetic analysis. This work was partially supported by grants from the Teikyo Heisei University (formerly Teikyo University of Technology).

Correspondence and requests for materials should be addressed to R.H. (e-mail: [renhym@ab.mbn.or.jp](mailto:renhym@ab.mbn.or.jp)).

## Ecological importance of the Southern Boundary of the Antarctic Circumpolar Current

Cynthia T. Tynan

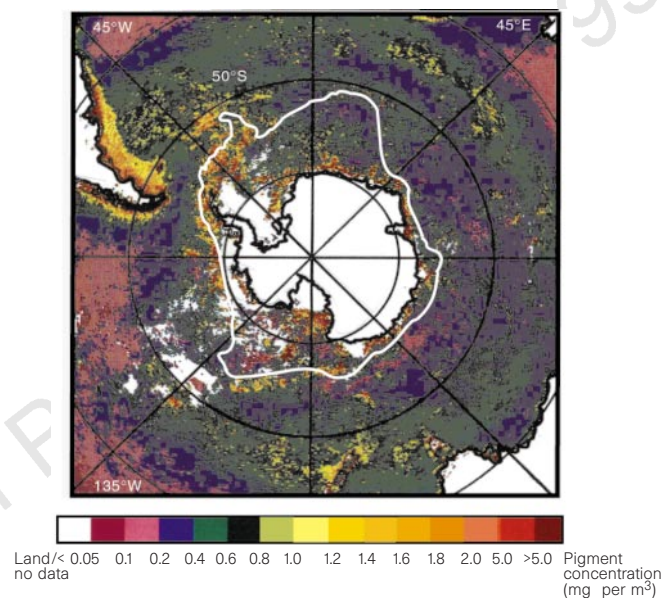
Research Associate of the National Research Council, National Marine Mammal Laboratory, NOAA, 7600 Sand Point Way NE, Seattle, Washington 98115, USA

The Southern Ocean surrounds the Antarctic continent and supports one of the most productive marine ecosystems. Migratory and endemic species of whales, seals and birds benefit from the high biomass of their principal prey, krill (*Euphausia superba*) and cephalopods, in this area. Most species of baleen whales and male sperm whales in the Southern Hemisphere migrate between low-latitude breeding grounds in winter and highly productive Antarctic feeding grounds in summer. Here I show the importance of the southernmost reaches of the strongest ocean current, the Antarctic Circumpolar Current (ACC), to a complex and predictable food web of the Southern Ocean. The circumpolar distributions of blue, fin and humpback whales from spring to midsummer trace the non-uniform high-latitude penetration of shoaled, nutrient-rich Upper Circumpolar Deep Water, which is carried eastward by the ACC. The poleward extent of this water mass delineates the Southern Boundary<sup>1</sup> of the ACC and corresponds not only to the circumpolar distributions of baleen whales, but also to distributions of krill and to regions of high, seasonally averaged, phytoplankton biomass. Sperm whales, which feed on cephalopods<sup>2</sup>, also congregate in highest densities near the Southern Boundary. The association of primary production, Krill, and whales with the Southern Boundary, suggests that it provides predictably productive foraging for many species, and is of critical importance to the function of the Southern Ocean ecosystem.

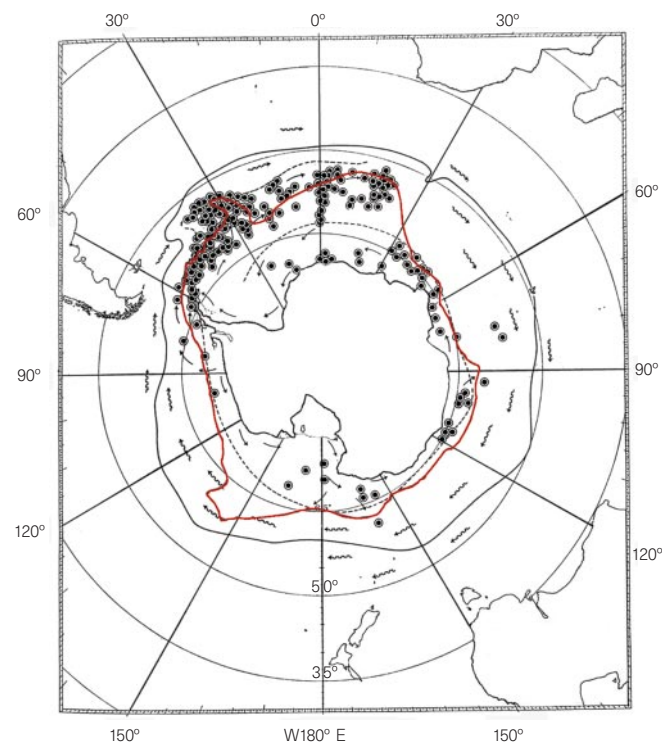
The ACC is dominated by a thick layer of warm, saline, oxygen-poor and macronutrient-rich water, Circumpolar Deep Water (CDW), which originates at low latitudes, shoals southwards, and is eventually entrained in the surface mixed layer<sup>1,3</sup>. Upper Circumpolar Deep Water (UCDW)<sup>1,3</sup> is the only water mass that is found exclusively in the ACC<sup>1</sup>. At its poleward edge, UCDW lies at depths near 200–500 m, where wind-induced divergence and upwelling bring high concentrations of phosphate, nitrate, and silicate to the Antarctic Surface Water<sup>3</sup>. The southern end of the characteristic signal of UCDW is a water-mass boundary which constitutes a reasonable poleward extent of the ACC; this feature has been named the Southern Boundary<sup>1</sup>.

The upward flux and entrainment of high concentrations of

macronutrients or trace metals into the surface mixed layer from UCDW at the Southern Boundary should increase regional primary production, enhance secondary production (that is, production of krill) and attract cetaceans and other foraging apex predators. In general, nutrients are not considered to be limiting for phytoplankton growth in Antarctic waters<sup>4</sup> and local iron enrichment may be an important factor that determines the spatial variability in phytoplankton production<sup>5</sup>. However, the concentrations of iron and other trace elements in UCDW and near the Southern Boundary



**Figure 1** Annual coastal zone colour scanner (CZCS) pigment concentration, averaged from data from October 1978 to June 1986 (ref. 4), in relation to the location of the Southern Boundary of the ACC<sup>1</sup> (white line).



**Figure 2** Distribution of principal concentrations of krill<sup>6</sup> in relation to the East Wind Drift and Weddell Drift (dashed lines), Polar Front (black line) and Southern Boundary<sup>1</sup> (red line). Principal concentrations of krill are shown as encircled black circles.

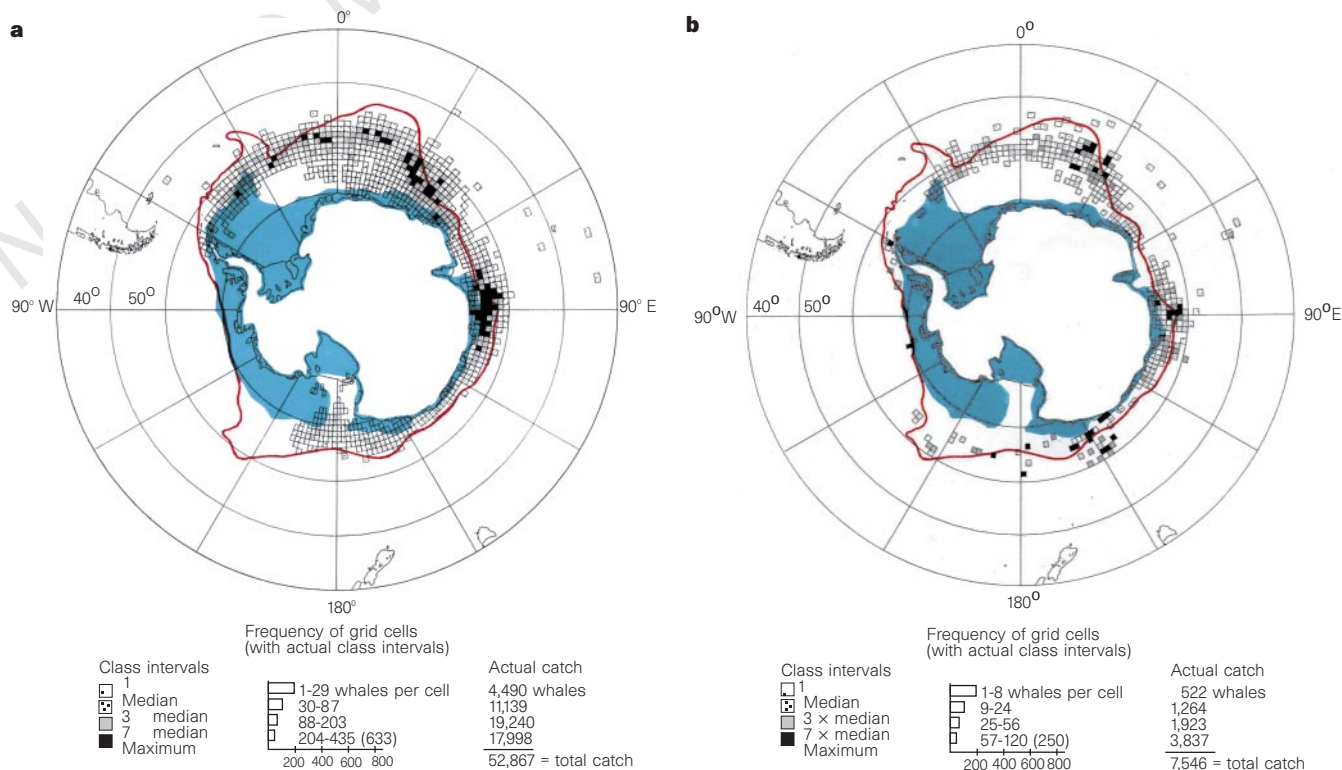
have not been examined. The Antarctic circumpolar distribution of phytoplankton pigment<sup>4</sup> shows that regions of higher primary productivity coincide with the Southern Boundary (Fig. 1). Where the Southern Boundary occurs close to the Antarctic continent (Southern Indian Ocean, 40°–150° E, and the southeast Pacific Ocean, 110° W to the Drake Passage), the highest phytoplankton pigment is confined to a narrower band along the continent. Conversely, where topography steers the Southern Boundary northwards away from the continent (east of the Antarctic Peninsula along the Weddell–Scotia Confluence, east of the Adeline Coast and north of the Ross Sea, and along the eastern flank of the Kerguelen Plateau), zones of higher productivity extend to much lower latitudes (Fig. 1).

The asymmetrical circumpolar distribution of krill<sup>6</sup> also matches the asymmetry in the poleward extent of the Southern Boundary between ocean basins (Fig. 2). Despite regional-scale temporal and spatial variability in the recruitment and abundance of krill<sup>7,8</sup>, the large-scale circumpolar pattern of principal krill concentrations still aligns with the Southern Boundary (Fig. 2). The duration and extent of ice cover has been linked to the spawning and recruitment success of krill<sup>7</sup>. Therefore, linkages between the Southern Boundary and the circumpolar pattern of sea ice should have a synergistic effect on krill distribution. Although the distribution of krill has been related to the prevailing surface currents<sup>6</sup>, the link between krill distribution and the Southern Boundary suggests the influence of shoaled UCDW. Eggs of krill are released at the surface, sink to deeper levels of warmer CDW, and hatch at depth<sup>9</sup>. The reproductive strategy of krill has evolved in response to CDW and the thermal structure, which enhances larval survival<sup>10</sup>. Therefore, the Southern Boundary should have an important effect on the distribution, development and feeding success of krill.

The broad, circumpolar distributions of whales from spring to midsummer also reflect the non-uniform high-latitude penetration of UCDW. Whales can occur as far south as the retreating ice edge<sup>11,12</sup>; however, the highest concentrations of blue whales

(*Balaenoptera musculus*), humpback whales (*Megaptera novaeangliae*) (Fig. 3) and fin whales (*Balaenoptera physalus*) coincide with the Southern Boundary in late spring and early summer. In December, the edge of the receding seasonal ice is near the Southern Boundary and the distribution of these three species of baleen whales aligns with both of these features. By January, however, when the ice has retreated closer to the continent, the highest densities of whales in some sectors of the Southern Ocean remain in ice-free areas near the Southern Boundary (Fig. 3). This may reflect processes at the Southern boundary as well as the development of the productive marginal ice zone<sup>13</sup> as the ice recedes. This distribution of the whales is particularly apparent in the South Atlantic from 30° W to 40° E, near the Kerguelen Plateau at 90° E, and in the South Pacific between 150° E and 150° W in regions where the Southern Boundary is further from the continent. Between the Weddell and Ross Gyres, where the Southern Boundary and the receding ice edge are closer to the Antarctic continent, higher densities of these three species of baleen whales also occur closer to the continent (40°–160° E; 60°–120° W). Past the eastern ends of the subpolar cyclones, near 30° E (ref. 14) and 130° W (ref. 15), the ACC shifts southward, bringing CDW close to the Antarctic Shelf Waters<sup>1</sup>. The distributions of whales during December and January also show this pattern of higher-latitude penetration in the regions between the Weddell and Ross Gyres (Fig. 3); presumably the whales are responding to the distribution of krill (Fig. 2) and the proximity of the ACC to the continental slope. This pattern shows that the Southern Boundary serves as a predictably productive foraging location for several months during the southward migration of the whales. As the pack ice retreats during the summer, with the continued development of the productive marginal ice-edge zone<sup>13</sup>, whales penetrate south of the Southern Boundary, approaching closer to the continent by February and March.

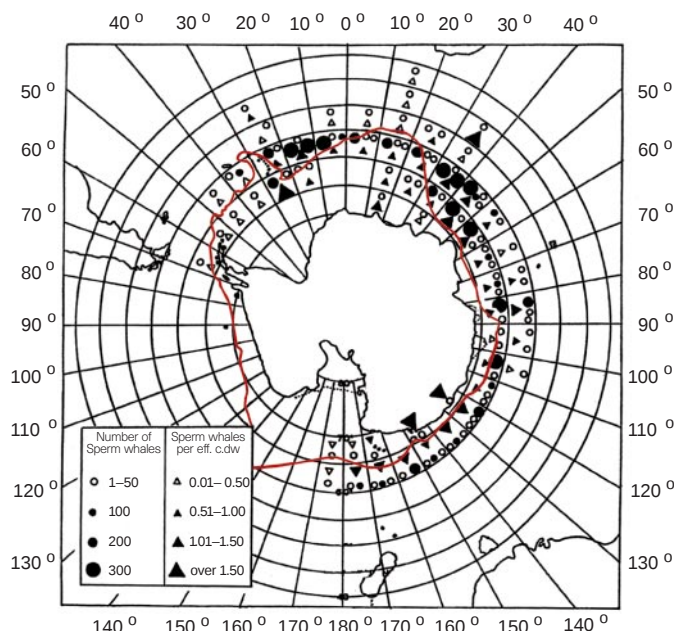
The circumpolar distribution of sperm whales (*Physeter macrocephalus*) also follows the Southern Boundary (Fig. 4). Sperm whales concentrate at higher latitudes in the Indian Ocean than in



**Figure 3** Distribution of whale catches. **a**, Blue whale, and **b**, humpback whale catches during January of 1931/1932–1966/1967 and 1931/1932–1962/1963, respectively<sup>17</sup>, in relation to the Southern Boundary<sup>1</sup> (red line) and the mean

monthly extent of sea-ice coverage (blue) for January of 1979–1987 (ref. 19). Grid size of the whale data is 1° latitude × 2° longitude.





**Figure 4** Distribution of sperm whale catches during the 1950s (ref. 18) in relation to the Southern Boundary<sup>1</sup> (red line). Catch data are gridded 5° latitude × 10° longitude, as total numbers (circles) and total catch per effective catcher's day's work (triangles).

the South Atlantic Ocean and track the increasing southward penetration of the Southern Boundary between 20°–60°E. Where the Southern Boundary occurs close to the Antarctic continent, such as in the Southern Indian Ocean, higher densities of sperm whales were historically found closer to the continent (Fig. 4). Regions in which sperm whales occurred in greater numbers in the 1950s, such as the South Sandwich Trench and northwest of Enderby Land, lie along or north of the Southern Boundary (Fig. 4). Therefore, male sperm whales seem to migrate southward as far as the poleward extent of UCDW. This suggests that the shoaling of UCDW may also affect the vertical movements and availability of squid, the preferred prey of sperm whales.

The presence of balaenopterids and sperm whales at the Southern Boundary indicates that processes there may enhance the availability of both krill and cephalopods. This suggests a trophic structure at the Southern Boundary that is different from the open-ocean trophic system of the Polar Frontal Zone in the Scotia Sea, where krill are absent and fishes are replaced by squid<sup>16</sup>. Pinnipeds and birds should also benefit from this complex and predictably productive feature. The Southern Boundary is identified here as a critical trophic structure in the function of the Southern Ocean ecosystem, determining the broad circumpolar distribution of productivity and the resulting cascade of trophic dynamics. □

## Methods

Monthly catch data from whaling operations in the Southern Ocean from 1931 to 1962 (ref. 17) were used to delineate the historical monthly distributions of blue, fin and humpback whales. Norwegian catch data of sperm whales from four seasons during the 1950s (ref. 18) were also analysed together with details of the Southern Boundary<sup>1</sup>. Mean monthly coverage of the Antarctic seasonal sea ice was obtained from satellite passive-microwave analyses for 1979–1987 (ref. 19). The whale-distribution data and the sea-ice data are not contemporary, as there are no circumpolar, synoptic, seasonal sea-ice data available for the whaling period 1931–1962.

Received 15 October 1996; accepted 15 January 1998.

1. Orsi, A. H., Whitworth, T. III & Nowlin, W. D. Jr On the meridional extent and fronts of the Antarctic Circumpolar Current. *Deep-Sea Res.* **142**, 641–673 (1995).

- Clarke, M. R. Cephalopoda in the diet of sperm whales of the Southern Hemisphere and their bearing on sperm whale biology. *Discov. Rep.* **37**, 1–324 (1980).
- Sievers, H. A. & Nowlin, W. D. Jr The stratification and water masses at Drake Passage. *J. Geophys. Res.* **89**, 10489–10514 (1984).
- Comiso, J. C. et al. Coastal zone color scanner pigment concentrations in the Southern Ocean and relationships to geophysical surface features. *J. Geophys. Res.* **98**, 2419–2451 (1993).
- de Baar, H. J. W. et al. Importance of iron for plankton blooms and carbon dioxide drawdown in the Southern Ocean. *Nature* **373**, 412–415 (1995).
- Marr, J. W. S. The natural history and geography of the Antarctic krill (*Euphausia superba* Dana). *Discov. Rep.* **32**, 33–464 (1962).
- Siegel, V. & Loeb, V. Recruitment of Antarctic krill *Euphausia superba* and possible causes for its variability. *Mar. Ecol. Prog. Ser.* **123**, 45–56 (1995).
- Priddle, J. et al. in *Antarctic Ocean and Resources Variability* (ed. Sahrhage, D.) 169–182 (Springer, Berlin, 1988).
- Laws, R. M. The ecology of the Southern Ocean. *Am. Sci.* **73**, 26–40 (1985).
- Hofmann, E. E. et al. Models of the early life history of *Euphausia superba*—Part I. Time and temperature dependence during the descent-ascend cycle. *Deep-Sea Res.* **39**, 1177–1200 (1992).
- Ribic, C. A., Ainley, D. G. & Fraser, W. R. Habitat selection by marine mammals in the marginal ice zone. *Ant. Sci.* **3**, 181–186 (1991).
- de la Mare, W. K. Abrupt mid-twentieth-century decline in Antarctic sea-ice extent from whaling records. *Nature* **389**, 57–59 (1997).
- Smith, W. O. Jr & Nelson, D. M. Importance of ice edge phytoplankton production in the Southern Ocean. *BioScience* **36**, 251–257 (1986).
- Orsi, A. H., Nowlin, W. D. Jr & Whitworth, T. III On the circulation and stratification of the Weddell Gyre. *Deep-Sea Res.* **140**, 169–203 (1993).
- Locarnini, R. A. *Water Masses and Circulation in the Ross Gyre and Environs* (thesis, Texas A&M Univ. 1994).
- Rodhouse, P. G. & White, M. G. Cephalopods occupy the ecological niche of epipelagic fish in the Antarctic Polar Frontal Zone. *Biol. Bull.* **189**, 77–80 (1995).
- Mizroch, S. A., Rice, D. W. & Larson, S. W. *Distribution of Rorquals in the Southern Ocean: An Atlas based on Pelagic Catch Data* (Fishery Bull., in the press).
- Holm, J. L. & Jørgensen, A. Occurrence of the sperm whale in the Antarctic and the possible influence of the moon. *The Norwegian Whaling Gazette* **4**, 161–182 (1959).
- Gloersen, P. et al. *Arctic and Antarctic Sea Ice, 1978–1987: Satellite Passive-Microwave Observations and Analysis* (NASA SP-511) (NASA, Washington DC, 1992).

**Acknowledgements.** I thank D. P. DeMaster, R. L. Gentry, G. L. Kooyman, D. G. M. Miller, A. H. Orsi and J. Priddle for review of the manuscript, and A. H. Orsi for discussions and for providing the coordinates of the Southern Boundary of the ACC.

Correspondence and requests for materials should be addressed to C.T.T. (e-mail: tynan@afsc.noaa.gov).

## Multiple stored views and landmark guidance in ants

S. P. D. Judd & T. S. Collett

Sussex Centre for Neuroscience, School of Biological Sciences, Brighton BN1 9QG, UK

Under some circumstances, Diptera and Hymenoptera learn visual shapes retinotopically, so that they only recognize the shape when it is viewed by the same region of retina that was exposed to it during learning<sup>1,2</sup>. One use of such retinotopically stored views is in guiding an insect's path to a familiar site<sup>3–5</sup>. Because the retinal image of an object changes with viewing distance and (sometimes) direction, a single stored view may be insufficient to guide an insect from start to goal. Little, however, is known about the number of views that insects store. Here we show that wood ants take several 'snapshots' of a familiar beacon from different vantage points. An ant leaving a newly discovered food source at the base of a landmark performs a tortuous walk back to its nest during which it periodically turns back and faces the landmark. The ant, on revisiting the familiar landmark, holds the edges of the landmark's image steady at several discrete positions on its retina. These preferred retinal positions tend to match the positions of landmark edges that the ant captured during its preceding 'learning walks'.

To investigate the possibility that wood ants (*Formica rufa*) store multiple views, we analysed their behaviour as they approached a cone. Our first step was to demonstrate that ants learn to distinguish between upright and inverted cones. An ant that had been trained to find sucrose close to either an upright or an inverted cone was presented with a choice between the two cones placed ~12 cm apart and ~40 cm from the ant (Fig. 1a). The ant approached the pair veering from one to the other. At a mean distance of 15.8 cm (s.d. = 7.2 cm,  $n = 76$ ) from the cones, it tended to select the



rewarded cone type and to walk to its base (Fig. 1b, c). Mostly, the ant aligned its body axis so that it faced the cone, suggesting that any views of the cone that the ant may have recorded were 'fixed' to the front of the eye.

Evidence for the storage of multiple views of the rewarded cone comes from a fine-grain analysis of the approaches of individual ants towards cones and towards a single black-and-white edge that was much longer than the border of the cone but oriented at the same 15° angle from the vertical. Let us start with edges appropriate to upright cones (Fig. 2a, d). If ants match the image of the edge to a single stored template, they should adjust their orientation during their approach so that the edge is kept appropriately positioned on the template in a single preferred retinal position (Fig. 1d). But if there are multiple templates, the edge should have several preferred retinal positions, one for each template (Fig. 1e, f).

Over some segments of the approach, the image of the edge is held relatively steady on the retina, but during other segments the edge moves rapidly as the ant scans its environment (Fig. 2b, e). We plotted how long the edge dwelt in different retinal positions during the plateau periods (a plateau is defined as a region of the plot where the edge remained stationary on the retina within a window of  $\pm 3^\circ$  for a distance of  $>6$  mm of approach) (Fig. 2c, f). The resulting histograms are multi-peaked in a way that suggests that the ant takes several snapshots of the cone and that it does so when it fixates the cone frontally. First, the two edges (Fig. 2a, d), corresponding to the right and left sides of the cone from the ant's perspective, are viewed predominantly by the right and left retina, respectively. Second, the peaks in the two plots are distributed symmetrically about the midline.

Translated into viewing distance, the positions of these peaks correspond to the cone viewed from 4, 7 and 17 cm. Similar data from seven more ants also showed multiple peaks. Peaks were identified automatically and the distribution of interpeak intervals shown in Fig. 3a was assembled from the approaches of all the ants to edges. The average angular distance between peaks is  $13.32^\circ$  (s.d.  $\pm 4.86^\circ$ ), which is about three times the  $4^\circ$  interommatidial separation, as measured in *Cataglyphis*<sup>6</sup>.

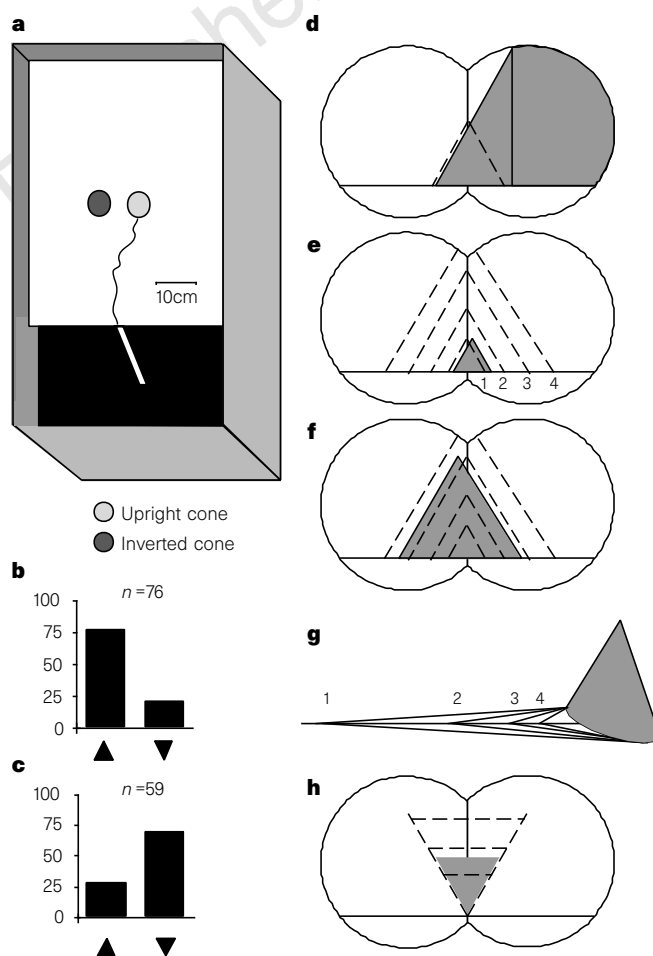
If these peaks are indeed evidence of multiple stored views, approaches to cones should generate peaks in similar retinal positions to approaches to black-and-white edges. With edges, each putative template can be matched throughout the approach by placing the edge in the appropriate position (Fig. 1d). However, with upright cones there is a unique matching position only at places where views are recorded (Fig. 1e, f). Away from those places, the cone is free to bounce between the two edges of the template. Nonetheless, the retinal plots of approaches to edges and cones (Fig. 2g, h) are similar. The two sets of peaks tend to coincide (compare Fig. 2c, f, i) and have similar distributions of interpeak intervals. To assess the statistical significance of this similarity, we measured the intervals between the closest peaks in the edge and cone distributions from the same ant (Fig. 3c). A Monte Carlo method indicates that the two sets of peaks obtained from each ant are aligned more closely than would be expected by chance ( $P < 0.002$  for data from eight ants).

A good check of our interpretation of multiple peaks is to examine where ants that are accustomed to approaching inverted cones position edges on their retina (Fig. 4). As the ant's horizon is only just above the ground plane, the apex of the inverted cone will be imaged at the same vertical position at all distances so that the bottom part of all the stored views will overlaid each other (Fig. 1h). Consequently, there should be only one preferred position for an edge. The pattern of results matches this prediction: the distribution of dwell-times for approaches to inverted cones and to related edges shows a single peak, which is centred on the midline (Fig. 4c, f).

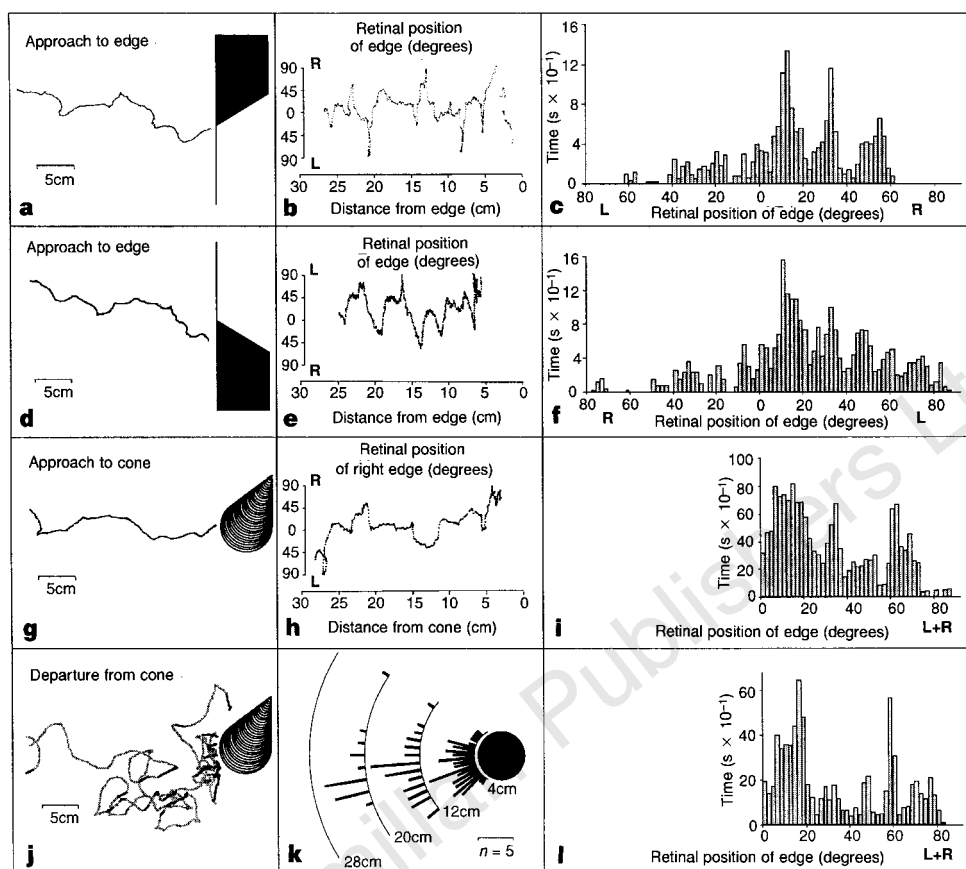
Further proof that ants trained to approach upright and inverted cones put the edges of these cones in different learnt retinal positions came from recording approaches of ants to a 1-cm-wide

black bar oriented at  $15^\circ$  to the vertical. Ants trained to approach inverted cones held the bottom of the bar on the midline as though they were centering the apex of the cone (Fig. 4i), whereas ants trained to approach upright cones kept the bottom of the bar positioned away from the midline (Fig. 4l), as would be appropriate when fixating an upright cone.

How does an ant approaching a familiar cone choose a template from its collection? If it uses the template that best matches its current view, it should engage small templates when far from the cone and larger ones close to. However, this mechanism of template selection would not necessarily lead to the same choice when the ant approaches a single extended edge: extended edges will, for example, match large templates equally well over the whole approach. In Fig. 5, we plot the mean retinal positions of the edge of the cone and of the extended edge as a function of the ant's distance from these edges. In both cases, the edges are, on average, relatively close to the midline when the ant is at the start of its journey and they move

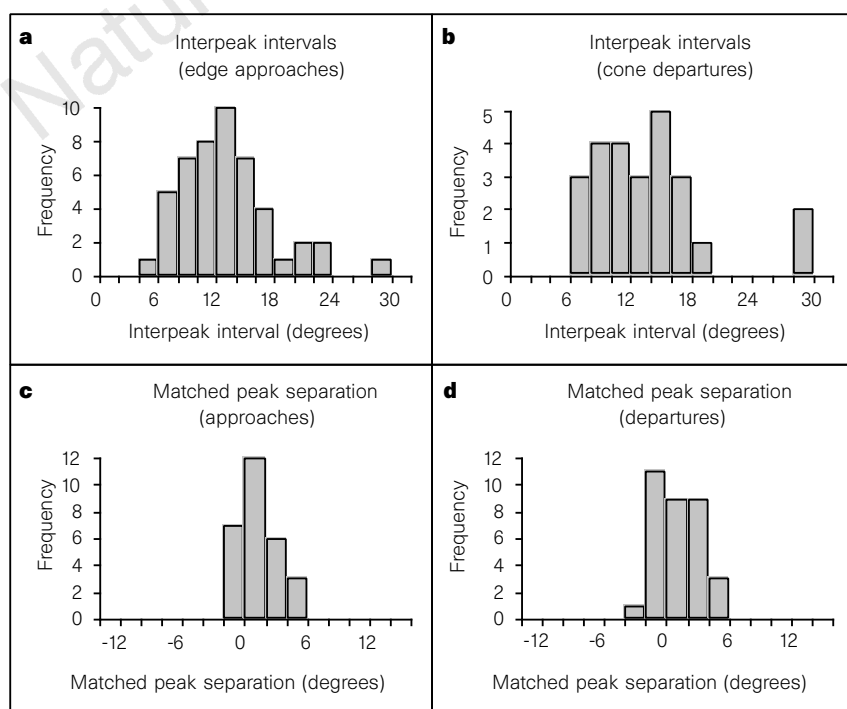


**Figure 1** Recognizing cones. **a**, Aerial view of ant nest and foraging area. The line going from ladder to cone is a typical example of an ant's path during a test. **b**, **c**, Histograms showing choice between the rewarded and unrewarded cones (in percent). Olfactory cues were eliminated by covering the shelf with fresh paper and using different cones in training and testing. A trial was ended and a choice scored when the ant approached within 3.5 cm of the centre of one or other cone. Statistical significance was assessed with the binomial test (**b**, upright cone rewarded,  $P < 0.001$ . **c**, Inverted cone rewarded  $P < 0.005$ ). **d**–**h**, Diagrams of the ant's visual field, with triangles representing hypothetical retinotopic templates of a cone. **d**, A single edge (grey shape) matched to a stored view of an upright cone. **e**, **f**, Four equally spaced retinotopic views of the cone, stored from different vantage points as indicated in **g**. **e**, **f**, Grey triangles represent the cone seen from a distance (**e**) and close up (**f**). **h**, Stored views of an inverted cone.

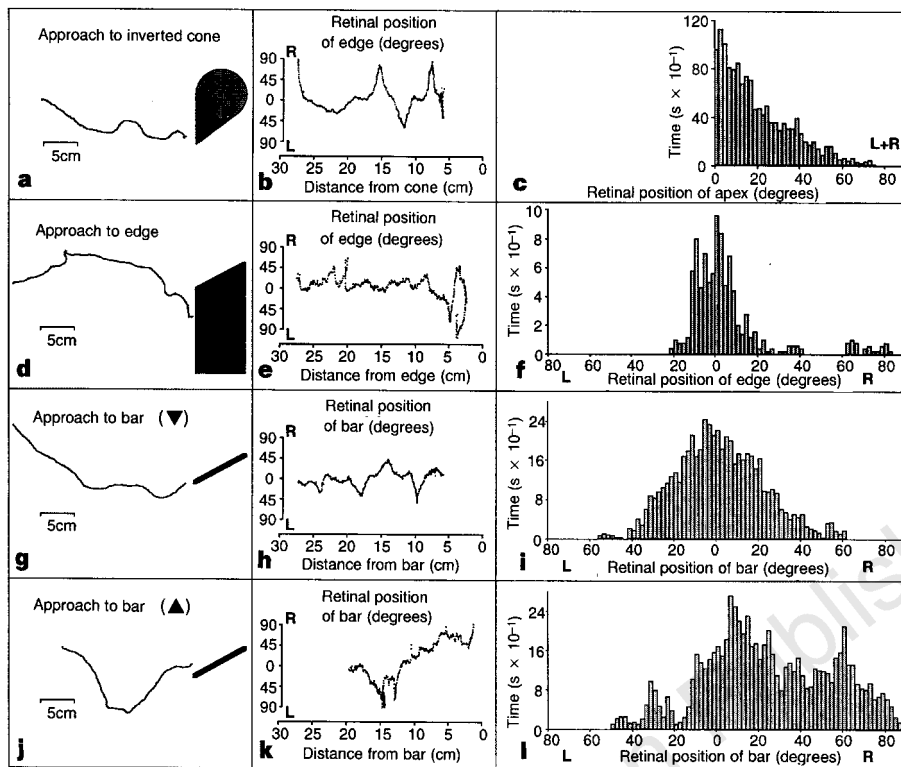


**Figure 2** A wood ant's approaches to extended black-and-white edges and a cone, and its return path from the cone. **a**, Trajectory towards a black-and-white edge that corresponds to the right edge of the cone (right side). **b**, Estimate position of right edge on the right retina is plotted against the ant's distance from the edge. R, right; L, left. **c**, Dwell time of the right edge in different retinal positions during the plateau phases of four approaches to the edge. Angles here and elsewhere are measured with respect to the bottom of the edge. Bins are 2° wide. **d**, Trajectory of an ant towards a black-and-white edge corresponding to the left edge of the cone. **e**, Estimated position of edge on the left retina during approach to left edge. **f**, Dwell time of edge in different positions on the retina during the plateau phases of five

approaches to left edge. Left and right are reversed to allow easier comparison between histograms. **g**, Approach of the same ant to a cone. **h**, Estimated position of one of the edges of the cone on the retina during the approach. **i**, Dwell time of both edges in different positions on the retina during the plateau phases of 14 approaches to the cone. Left and right sides of the retina are superimposed. **j**, Trajectory of same ant during its first departure from the cone. The trace is darkened during inspections (that is, when the ant turns and walks back towards the cone). **k**, Distribution of inspections of the cone accumulated over 14 departures. **l**, Dwell time of both edges of the cone in different retinal positions during the inspections from 14 departures. Left and right sides of the retina are superimposed.



**Figure 3** Distribution of peaks. **a**, Histogram of the intervals between neighbouring peaks in multimodal distributions derived from approaches to extended edges such as those used for Fig. 2c, f. Data from eight ants. **b**, As for **a**, but for departures from cones; data from seven ants. **c**, Histogram of the differences between the closest peaks from paired multimodal distributions of approaches to extended edges and approaches to cones. Data from eight ants. **d**, As for **c**, but for differences between approaches to extended edges and departures from cones. Data are from seven ants.



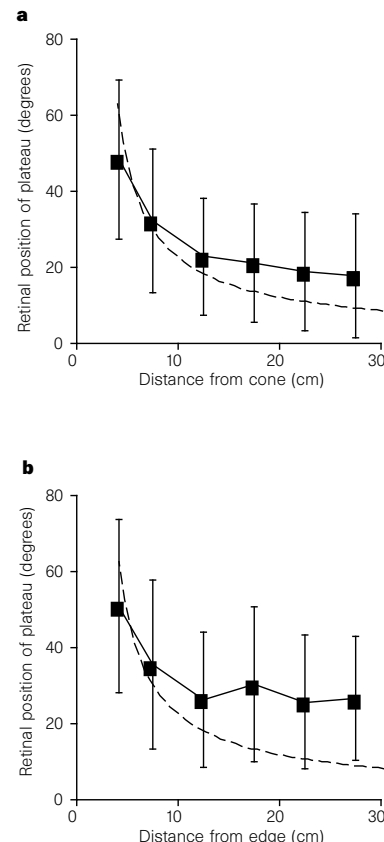
**Figure 4** Approaches to inverted cones and black bars. **a–d**, Approaches of an ant trained to approach an inverted cone. **a**, Approach of ant to an inverted cone. **b**, Trajectory of the edge of the cone over the retina. R, right; L, left. **c**, Dwell time of the edge in different retinal positions during the plateau phases of 20 approaches to the cone. **d**, Approach of an ant to an extended edge. **e**, Trajectory of the edge over the retina. **f**, Dwell time of edge in different retinal positions during the plateau phases of six approaches to the edge. Similar data were recorded from four further ants trained to approach inverted cones. **g–l**, The distribution of retinal positions of a 1-cm wide, black bar oriented at 15° to the vertical, accumulated over 46 approaches of ants trained to approach an inverted cone. Angles are measured with respect to the midpoint of the bottom of the bar (**g**) and 30 approaches of ants trained to approach an upright cone (**j**). Data were included when the ant was between 0 cm and 12 cm from the bar. To show that the two conditions induced significantly different approaches, the mean retinal position was calculated for each approach and the two sets of means compared. Probability that approaches come from the same distribution is approximately 0.01 (Wilcoxon test). **h, k**, Trajectory of the edge of the bar over the retina. **i, l**, Dwell time of the bar in different retinal positions during the plateau phase.

towards the periphery of the retina as the ant nears the stimulus. The ant tends to select a template suited to a distant cone when it views either the edge or the cone at a distance, and to select a template suited to a close cone when it is near. The similarity of the two plots indicates that cues in addition to static-image matching may be involved in template selection.

Bees and wasps perform elaborate learning flights on first leaving their nest or a newly discovered feeding place. During these flights they turn back and look at the goal<sup>7–10</sup>. Wood ants do something similar. On their first few departures from the cone, their path is very tortuous, with much turning back. The darkened line in the departure of Fig. 2j indicates path segments in which the ant has turned back and is walking towards the landmark. When the ant is close to the landmark, the retinal image transforms rapidly and these inspections of the cone are frequent and occur over a large range of viewing directions, perhaps helping the ant separate the landmark from the background. Further away, the ant's view of the landmark changes more gradually; inspections become correspondingly rarer and the range of viewing directions is narrower (Fig. 2j, k).

One indication that learning may occur during inspections of the cone on homeward journeys is the distribution of retinal positions of the edge of the cone when the ant looks back at it. Peaks occur in similar locations to the comparable peaks derived from the same ant's approaches to edges. A Monte Carlo method shows that the intervals between the closest matching peaks on approaches and departures (Fig. 3d) are smaller than would be expected by chance ( $P < 0.004$  for data from seven ants). Some of the peaks found on approaches may be missing from the distribution of peaks generated during departures (compare Fig. 2a, l). The absence of some peaks indicates that, as in bees<sup>9</sup>, the acquisition of landmark information cannot be restricted to departures.

Our results indicate that object recognition by insects and humans may have interesting parallels. In both cases, there is evidence<sup>11–15</sup> that the visual system encodes familiar objects in terms of stored two-dimensional views and that the problem of recognizing the same object from several viewpoints is solved, in



**Figure 5** Retinal position of the edge of an upright cone or of a black-and-white edge is plotted against the ant's distance from: **a**, an upright cone (66 approaches); and **b**, a black-and-white edge (20 approaches). The plots average data from all approaches of all ants so that no preferred retinal positions will be apparent. Error bars show  $\pm 1$  s.d. The dashed curve represents the position of the cone edge when assuming that the cone is fixated centrally.



part, by acquiring several two-dimensional views. The plots in Fig. 2c, f and k indicate that ants may have adopted a strategy of sampling more often close to the cone, where image size changes rapidly, and less frequently further away, where image size is less dependent on range. Sampling the cone at equi-angular separations gives the ant a series of views that evenly covers the transformation of the shape from its first sighting to the ant's arrival at its base (Fig. 1e, f, g). □

## Methods

**Housing, training and testing ants.** Ant nests were kept in large plastic tanks. Individually marked foragers collected sucrose on a flat shelf (120 cm by 80 cm) above the nest which they reached by means of a paper 'drawbridge' (Fig. 1a). Ants from one colony foraged for sucrose at the base of a black upright cone (7-cm base, 12-cm height). The cone was inverted for ants from another colony. Both cone types were present during the training for the choice experiment (Fig. 1) and the positions of the cones were swapped frequently. One cone was associated with the sucrose and the other was an unrewarded distractor. The shelf was wiped with alcohol after each foraging trip to remove possible trail pheromones.

**Videorecording and analysis.** The ~30-cm trips to and from the feeder were videorecorded from above and analysed automatically to obtain the ant's position and the orientation of its body axis. The retinal position of the edge of the cone was calculated on the assumption that the head is aligned with the body. High-magnification videorecordings of ants as they walked straight and turned corners showed that the head is kept mostly aligned with the body. On straight paths, the mean angle between head and body was  $0.013^\circ$  (s.d. =  $\pm 1.99^\circ$ ,  $n = 90$  frames), and while the ant was performing a  $83^\circ$  turn the mean angle between head and body was  $2.57^\circ$  (s.d. =  $\pm 3.78^\circ$ ,  $n = 23$  frames).

Plateaux were selected using a computer algorithm that fitted straight-line segments to the plots of the retinal position of edges. The plateaux collected from all the approaches made by one ant to the same stimulus were sorted into a histogram, which displays how long the edge was held in different retinal positions. Peaks in the histogram were then found using another computer algorithm. The data were first smoothed with a binomial filter of width 1, and the peak positions were given by the positions of the remaining local maxima.

**Assessing the similarity of multimodal distributions.** Statistical tests were performed on the data from each ant using the positions of the peaks. We compared the similarity of peak positions from approaches to cones and approaches to extended edges, or from departures from cones and approaches to extended edges. The computer picked out the closest matching pairs of peaks in the two distributions as follows: for any given peak in the edge-derived data, the closest peak in the cone-derived data was found. Provided that the reverse also held, that is, that the original peak in the edge-derived data was the closest peak to the matched peak in the cone-derived data, the interpeak difference was recorded and the pair was removed from the data set. This was continued until there were no further matches. The process was repeated with data from the other ants, and the standard deviation of all of the peak differences, assuming a mean difference of zero, was taken as a measure of the similarity of the distributions. The better the match between the peak distributions, the smaller this similarity score was.

5,000 sets of artificially generated cone-derived data were then matched to the real edge-derived data using the same procedure. For each ant, the artificial data sets each had the same number of peaks as the real data, but the peaks were distributed randomly between  $0^\circ$  and  $80^\circ$  from the midline, with the constraint of a minimum separation of  $6^\circ$  and a maximum separation of  $24^\circ$  between peaks. We computed a similarity score across all ants for each data set. The number of sets out of the 5,000 that scored less than the real data was taken as the probability of getting the real match by chance<sup>16</sup>.

Received 3 December 1997; accepted 19 February 1998.

1. Dill, M., Wolf, R. & Heisenberg, M. Visual pattern recognition in *Drosophila* involves retinotopic matching. *Nature* **365**, 751–753 (1993).
2. Wehner, R. in *Handbook of Sensory Physiology* Vol. VII/6C (ed. Autrum, H.) 287–616 (Springer, Berlin, Heidelberg, New York, 1981).
3. Collett, T. S. & Land, M. F. Visual spatial memory in a hoverfly. *J. Comp. Physiol.* **100**, 59–84 (1975).
4. Wehner, R. & R ber, F. Visual spatial memory in desert ants, *Cataglyphis fortis* (Hymenoptera, Formicidae). *Experientia* **35**, 1569–1571 (1979).

5. Cartwright, B. A. & Collett, T. S. Landmark learning in bees: experiments and models. *J. Comp. Physiol.* **151**, 521–543 (1983).
6. Zollikofer, C. P. E., Wehner, R. & Fukushima, T. Optical scaling in conspecific *Cataglyphis* ants. *J. Exp. Biol.* **198**, 1637–1646 (1995).
7. Vollbrecht, J. Zur Orientierung junger Honigbienen bei ihrem ersten Orientierungsflug. *Zool. Jb.* **79**, 33–69 (1975).
8. Zeil, J. Orientation flights of solitary wasps (*Cerceris*; Sphecidae; Hymenoptera): I. Description of flight. *J. Comp. Physiol.* **A172**, 189–205 (1993).
9. Lehrer, M. Why do bees turn back and look? *J. Comp. Physiol.* **A172**, 549–563 (1993).
10. Collett, T. S. & Lehrer, M. Looking and learning: a spatial pattern in the orientation flight of the wasp *Vespa vulgaris*. *Proc. R. Soc. Lond. B* **252**, 129–134 (1993).
11. Rock, I., DiVita, J. & Barbeito, R. The effect on form perception of change of orientation in the third dimension. *J. Exp. Psychol. Hum. Percept. Perform.* **7**, 719–732 (1981).
12. Rock, I. & DiVita, J. A case of viewer centred object perception. *Cogn. Psychol.* **19**, 280–293 (1987).
13. Tarr, M. J. & Pinker, S. Mental rotation and orientation-dependence in shape recognition. *Cogn. Psychol.* **21**, 233–282 (1989).
14. B lthoff, H. H. & Edelman, S. Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proc. Natl Acad. Sci. USA* **89**, 60–64 (1992).
15. Ullman, S. *High Level Vision: Object Recognition and Visual Cognition* (MIT Press, Cambridge, MA, 1996).
16. Manly, B. F. J. *Randomization and Monte Carlo Methods in Biology* (Chapman and Hall, London, 1991).

**Acknowledgements.** We thank M. F. Land and D. Osorio for valuable comments. Financial support came from the BBSRC and Human Frontier Science Program. S.P.D.J. received a BBSRC research studentship.

Correspondence and requests for materials should be addressed to S.P.D.J. or T.S.C.

## Encoding of three-dimensional structure-from-motion by primate area MT neurons

David C. Bradley, Grace C. Chang & Richard A. Andersen

Division of Biology, California Institute of Technology,  
391 South Holliston Avenue, Pasadena, California 91125, USA

We see the world as three-dimensional, but because the retinal image is flat, we must derive the third dimension, depth, from two-dimensional cues. Image movement provides one of the most potent cues for depth<sup>1–6</sup>. For example, the shadow of a contorted wire appears flat when the wire is stationary, but rotating the wire causes motion in the shadow, which suddenly appears three-dimensional. The neural mechanism of this effect, known as 'structure-from-motion', has not been discovered. Here we study cortical area MT, a primate region that is involved in visual motion perception. Two rhesus monkeys were trained to fixate their gaze while viewing two-dimensional projections of transparent, revolving cylinders. These stimuli appear to be three-dimensional, but the surface order perceived (front as opposed to back) tends to reverse spontaneously. These reversals occur because the stimulus does not specify which surface is in front or at the back. Monkeys reported which surface order they perceived after viewing the stimulus. In many of the neurons tested, there was a reproducible change in activity that coincided with reversals of the perceived surface order, even though the stimulus remained identical. This suggests that area MT has a basic role in structure-from-motion perception.

We trained two rhesus monkeys to fixate on a stationary target while we showed the two-dimensional projection of a revolving, random-dot cylinder (Fig. 1a). This projection contains opposite-going motions which convey a sense of front and back, or surface order. Monkeys then reported the direction of the front surface by glancing at one of two targets that appeared on either side of the cylinder's former position. As two-dimensional projections are flat, they do not specify the surface order, so the monkeys' answers reflect their three-dimensional perception of the stimulus.

We also showed rotating cylinders whose structure (and thus surface order) was specified with disparity (depth). Some of these were flattened by multiplying the depth of each dot by a fraction (0, 12.5, 25, 50 or 100%). All cylinders had their centre at zero disparity, so one surface appeared to be near, the other far, relative to the

fixation depth. Figure 1b shows that performance—the ability to judge surface order—decreased predictably as the disparity decreased, suggesting that monkeys were doing the task as required.

In simultaneous, single-neuron recordings of MT activity, we oriented cylinders so that one surface moved in the neuron's preferred direction (determined in preliminary tests), the other in the opposite direction. We expected responses to depend on surface order because MT cells tend to prefer motion either behind or in front of the fixation point (far or near)<sup>7</sup>. Preferred-direction motion on the 'active' side tends to excite, whereas antipreferred motion on the active side tends to suppress<sup>8</sup>. Therefore, one of the two surface orders should be optimal because it places preferred motion on the active side while placing antipreferred motion on the other side. Indeed, when the highest-disparity cylinders were shown, 68/109 MT cells responded significantly better to one of the two surface orders ( $P < 0.05$ ,  $t$ -test).

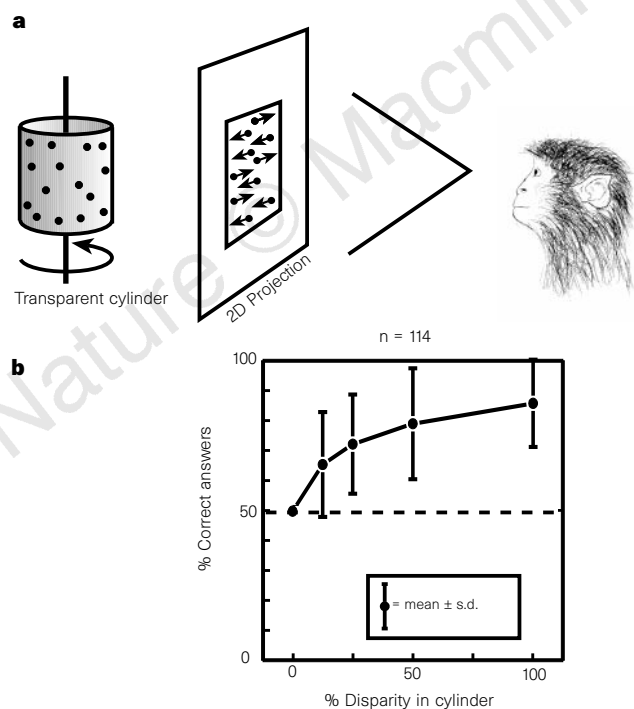
Responses were linked to the perception of surface order. The cell in Fig. 2, for example, preferred the front going down, the back up, and this was true whether that surface order was specified with disparity (top) or simply perceived as such in a zero-disparity stimulus (bottom). Of the 68 cells with a preferred surface order (see above), 34 responded differently when the stimulus was perceived with different surface orders ( $P < 0.05$ ,  $t$ -test; see Methods). Most cells (27/34) showed 'correlated' behaviour, meaning that responses were higher when the neuron's preferred order (defined at the highest disparity) was perceived. This was true for both types of cells irrespective of whether they had a preferred direction in front (17/20) or in the back (7/9; 5 cells not classifiable

as near or far). Given the low frequency of cells with the opposite behaviour (7 of 68 possible), it is not clear whether a distinct cell class of this type really exists.

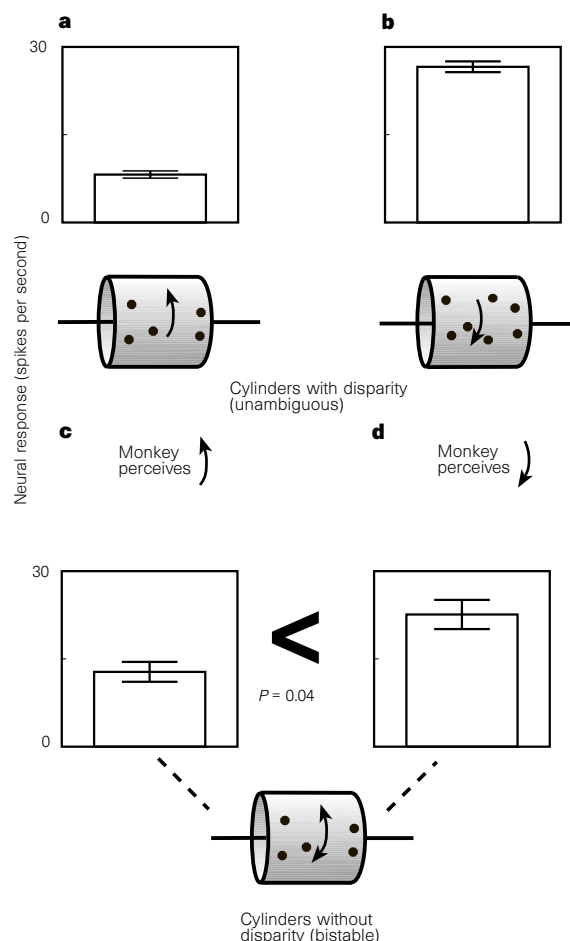
Although disparity cues bias perception in favour of a particular surface order, all stimuli were potentially bistable (could be perceived either way) (Fig. 1b). Figure 3a shows that whatever the disparity, and whatever the specified surface order, responses were higher when the neuron's preferred surface order was perceived (compared with the non-preferred order). Moreover, whether the variable was the specified surface order (100% disparity) or the perceived surface order (zero disparity), the time course over which activity diverged was similar (Fig. 3b). Thus, to the extent that the perceived surface order of a given stimulus differed, MT activities tended to reflect that difference.

Previous experiments with flat patterns showed that opposite motion directions suppress MT responses, but that this suppression decreases, and in some cases changes to facilitation, when opposite directions are shown at different disparities<sup>8,9</sup>. This suggests that there are inhibitory connections between MT cells tuned for opposite directions and similar depths (Fig. 4a), and excitatory connections between cells tuned for opposite directions and different depths (that is, near rather than far). Such depth-dependent interactions may be important for computing surface movement because they emphasize coherent (same direction) motion signals while suppressing random signals (motion noise) from a given surface<sup>8,10</sup>.

Structure-from-motion perception may begin with the bistable nature of this circuitry. MT cells typically prefer either near or far



**Figure 1** The monkeys' task and average performance. **a**, Monkeys viewed two-dimensional projections of three-dimensional, revolving cylinders, then reported the direction of rotation they perceived. **b**, Psychometric function (all data pooled and averaged) relating performance to the amount of disparity in the cylinders. The abscissa shows disparity as a percentage of the disparity that would be seen looking at a real (3D) cylinder. The ordinate shows the percentage of correct answers regarding the surface order. Error bars show standard deviation.



**Figure 2** Data from an MT neuron. **a**, **b**, Responses to 100%-disparity cylinders (using only trials with correct responses). **c**, **d**, Responses of the same cell to zero-disparity cylinders, perceived as going front-up (**c**) or front-down (**d**).

stimuli, but their tuning is broad enough that they also respond to zero-disparity stimuli<sup>7</sup>. Therefore, a zero-disparity cylinder projection could potentially activate four neuronal pools, tuned (assuming a vertical cylinder) for near-right, near-left, far-right and far-left (Fig. 4a). But because of the inhibition and excitation discussed above, an even distribution of activity would be unstable, tending to 'fall' into a distribution that places opposite directions in different depth channels. For example, an increase in the activity of near-right cells could lead to a suppression of near-left cells and an activation of far-left cells; the far-left cells, in turn, would suppress the far-right cells, and so on. The resulting activity distribution would be concentrated in the near-right and far-left channels, presumably resulting in the perception of the front surface moving right, the back moving left (Fig. 4b). On different trials, activity might instead end up in the near-left and far-right channels, depending on the adaptation, or fatigue, of the different channels at the outset of each trial<sup>11</sup>.

In some MT neurons, activity increases when the monkey pays attention to that neuron's preferred direction<sup>12</sup>. Assuming our monkeys always attended the cylinder's front surface, this could produce an artefact by increasing activity when the neuron's preferred direction appears in front. However, many neurons responded best when their preferred direction was in the back. Moreover, when a neuron preferred a given surface order (based on

disparity), it typically responded best when that order was perceived. This cannot be explained by an attention effect, unless the monkeys learned selectively to attend to one of the two surfaces, depending on the response properties of the neuron currently being tested. This is extremely unlikely.

Area MT is no doubt specialized for motion computation<sup>13,14</sup>, but there is accumulating evidence that it also has a role in three-dimensional surface representation. MT neurons have large receptive fields, capable of spatially integrating motion cues; they are direction- and depth-selective, consistent with surface-oriented motion computation; and they exhibit direction opponency, which may be used for surface-specific noise reduction<sup>7-10,15</sup>. MT neurons are thus well suited to the task of transforming motion cues into information about surfaces and depth. In fact, MT lesions impair monkeys in tasks where three-dimensional structure is judged from motion cues<sup>5,6</sup>. Up to now, however, there has been no direct evidence that the perception of structure is linked to activities in MT. Our findings provide this evidence, and as such they suggest that MT has a central role in structure-from-motion perception. Of course, perception may occur in a different area which receives input from MT. But wherever it occurs, our findings indicate that the perception of structure is ultimately influenced by the segregation of MT activity into separate depth channels. □

# Methods

**Stimuli.** Visual stimuli were shown on a 21-inch CRT screen, 57 cm from the monkey's eyes. Isolated neurons were tested to find their approximate receptive field, and subsequent stimuli were centred within this field. Neurons were tested for single-pattern direction selectivity as described<sup>8</sup>.

Cylinder projections were 7° wide and 7° tall, contained 150 randomly placed dots<sup>8</sup>, and rotated at 100° per second. Cylinders were positioned with their centre 3.0–8.2° from the fixation point, and with 10/109 exceptions, no part of the cylinder overlapped the fixation point. All cylinders had their centre at zero disparity, so one surface appeared near, the other far, relative to the fixation depth. Dots were rendered in stereo with an anaglyph system<sup>8</sup>.

Most neurons were tested with cylinders containing 5 disparity levels (0, 12.5, 25, 50 and 100%) as explained in the text. However, in preliminary tests with 31 of the neurons, only two disparities were tested: one low (0 or 12.5%) and one high (25–100%). Results from these cells were similar to those overall, so they were combined with the remaining cells to form the present data set.

**Task.** Monkeys fixated for 0.5 s before, 1 s during, and 0.5 s after the 1-s cylinder presentation. Selection targets were 5.5° on either side of the cylinder's rotation axis, positioned on a line bisecting the cylinder's height. Monkeys were rewarded with a drop of juice for choosing the target corresponding to the direction of the front surface (for zero-disparity cylinders, rewards were given randomly at a frequency of 80%).

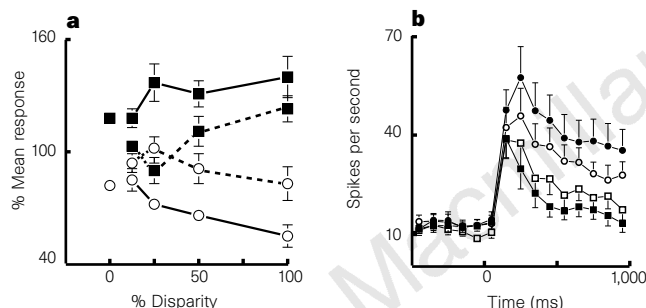
**Eye position.** Monkeys were required to fixate within a 3° square window while the cylinder appeared on the screen. Subsequent analysis showed that eye position remained inside a 1° window in 97% of the trials. The within-trial standard deviation of eye position, sampled at 100 Hz, was 0.05° horizontal, 0.12° vertical.

In one of the monkeys, both eye positions were measured simultaneously to calculate the depth of fixation (units of degrees angular disparity). Standard deviation was 0.05° within trials and 0.07° from trial-to-trial. Comparing trials in which opposite surface orders were perceived revealed no differences in fixation depth ( $P \geq 0.05$  in 95% of the  $t$ -tests;  $n = 53$ ).

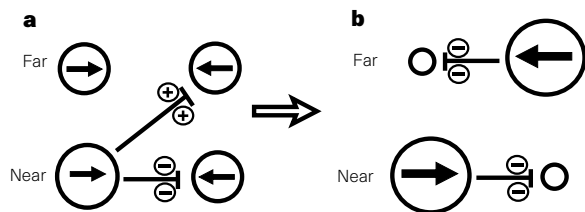
**Analysis.** Our main analysis involved testing for a response difference ( $P < 0.05$ ,  $t$ -test) associated with the perceived surface order of one or more of the stimuli (each stimulus defined by a given disparity and surface order). Multiple  $t$ -testing can in theory increase the false-positive rate, but it cannot account for the high percentage of 'correlated' cells (see text), because false positives have an even chance of being correlated.

Received 4 August 1997; accepted 20 January 1998.

1. Rogers, B. J. & Graham, M. E. Similarities between motion parallax and stereopsis in human depth perception. *Vision Res.* **22**, 261–270 (1982).
2. Miles, W. R. Movement interpretations of the silhouette of a revolving fan. *Am. J. Psychol.* **43**, 392–405 (1931).
3. Wallach, H. & O'Connell, D. N. The kinetic depth effect. *J. Exp. Psychol.* **45**, 205–217 (1953).



**Figure 3** Averaged data from 'correlated' MT cells ( $n = 27$ ). **a**, Perceptual effects at different disparities. Responses were normalized to the mean response for each disparity and plotted as a function of disparity. Data show mean responses  $\pm$  standard error. Top two curves (squares): actual surface order matches the neuron's preferred surface order. Lower two curves (circles): actual surface order is opposite to the neuron's preferred order. Solid lines, correct trials; broken lines, error trials. At zero disparity, as there is no actual surface order (the stimulus has zero disparity), the data points represent the two perceived surface orders. **b**, Time course of responses at zero disparity (inner traces) and 100% disparity (outer traces; correct trials only). Circles represent the neurons' preferred surface order (perceived or actual); squares represent the non-preferred order. The stimulus was visible from 0–1,000 ms. Data show mean response  $\pm$  standard error (in some cases error bars are smaller than the symbols).



**Figure 4** Proposed model to explain how suppression and facilitation in MT could give rise to the illusion of depth: **a**, a cylinder projection could activate four neuronal pools; **b**, because of excitatory and inhibitory interactions, activity migrates into opposite-direction, separate-depth channels. This presumably creates the perception of a cylinder rotating in a single direction. Circle diameter denotes magnitude of activity.



4. Ringach, D. L., Hawken, M. J. & Shapley, R. Binocular eye-movements caused by the perception of 3-dimensional structure-from-motion. *Vision Res.* **36**, 1479–1492 (1996).
5. Siegel, R. M. & Andersen, R. A. Perception of three-dimensional structure from motion in monkey and man. *Nature* **331**, 259–261 (1988).
6. Andersen, R. A. & Siegel, R. M. in *Signal and Sense: Local and Global Order in Perceptual Maps* (eds Edelman, G. M., Gall, W. E. & Cowan, W. M.) 163–184 (Wiley, New York, 1990).
7. Maunsell, J. H. & Van Essen, D. C. Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *J. Neurophysiol.* **49**, 1148–1167 (1983).
8. Bradley, D. C., Qian, N. & Andersen, R. A. Integration of motion and stereopsis in middle temporal cortical area of macaques. *Nature* **373**, 609–611 (1995).
9. Snowden, R. J., Treue, S., Erickson, R. G. & Andersen, R. A. The response of area MT and V1 neurons to transparent motion. *J. Neurosci.* **11**, 2768–2785 (1991).
10. Hildreth, E. C., Ando, H., Andersen, R. A. & Treue, S. Recovering three-dimensional structure from motion with surface reconstruction. *Vision Res.* **35**, 117–137 (1995).
11. Nawrot, M. & Blake, R. A neural network model of kinetic depth. *Visual Neurosci.* **6**, 219–227 (1991).
12. Treue, S. & Maunsell, J. H. R. Attentional modulation of visual-motion processing in cortical areas MT and MST. *Nature* **382**, 539–541 (1996).
13. Newsome, W. T., Britten, K. H. & Movshon, J. A. Neuronal correlates of a perceptual decision. *Nature* **341**, 52–54 (1989).
14. Logothetis, N. K. & Schall, J. D. Neuronal correlates of subjective visual perception. *Science* **245**, 761–763 (1989).
15. Bradick, O. Segmentation versus integration in visual-motion processing. *Trends Neurosci.* **16**, 263–268 (1993).

**Acknowledgements.** We thank B. Gillikin and S. Gertmenian for technical assistance and F. Crick and C. Koch for comments on the manuscript. Supported by grants from the National Eye Institute, the Human Frontier Science Program, and the Sloan Foundation for Theoretical Neurobiology.

Correspondence and requests for material should be addressed to R.A.A. (e-mail: andersen@vis.caltech.edu).

## Glycine-receptor activation is required for receptor clustering in spinal neurons

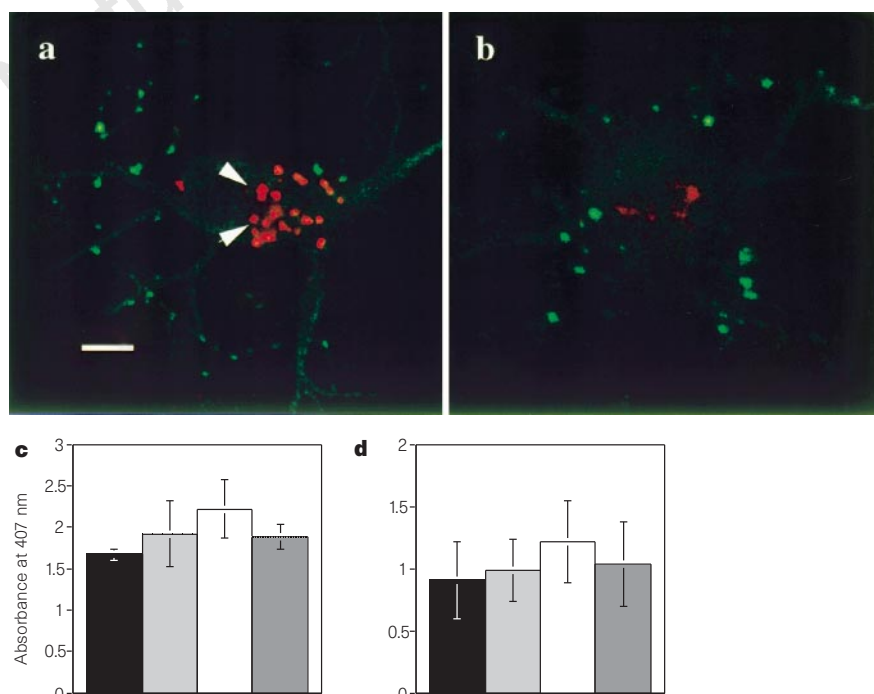
J. Kirsch & H. Betz

Department of Neurochemistry, Max-Planck-Institute for Brain Research, Deutschordenstrasse 46, 60528 Frankfurt/Main, Germany

The ability of nerve cells to receive up to several thousands of synaptic inputs from other neurons provides the anatomical basis for information processing in the vertebrate brain. The formation of functional synapses involves selective clustering of neurotransmitter receptors at presumptive postsynaptic regions of the neuronal plasma membrane<sup>1–4</sup>. Receptor-associated proteins are believed to be crucial for this process. In spinal neurons, synaptic

targeting of the inhibitory glycine receptor (GlyR)<sup>5,6</sup> depends on the expression of the anchoring protein gephyrin<sup>7–9</sup>. Here we show that the competitive GlyR antagonist strychnine and L-type  $\text{Ca}^{2+}$ -channel blockers inhibit the accumulation of GlyR and gephyrin at postsynaptic membrane areas in cultured rat spinal neurons. Our data are consistent with a model in which GlyR activation that results in  $\text{Ca}^{2+}$  influx is required for the clustering of gephyrin and GlyR at developing postsynaptic sites. Similar activity-driven mechanisms may be of general importance in synaptogenesis.

Inhibitory neurotransmission in adult spinal cord is largely mediated by the strychnine-sensitive glycine receptor GlyR<sup>10</sup>. The developmental sequence of GlyR expression is recapitulated in primary cultures prepared from rodent spinal cord at embryonic day 14 (refs 11, 12). After 6–10 days *in vitro*, more than 90% of the neurons in these cultures express functional glycine receptors<sup>9,13,14</sup>. By using double labelling with the GlyR-specific monoclonal antibody 4a (ref. 15) and a polyclonal antibody raised against synaptic vesicle proteins (anti-SVP)<sup>16</sup>, it has been shown that synaptic contacts between neurons that are characterized by the close apposition of pre- and postsynaptic antigens are established after one week in culture<sup>9,14</sup>. We find that early treatment (within 24 h of plating) of rat spinal cord cultures with 30  $\mu\text{M}$  tetrodotoxin, which blocks neuronal activity by inhibiting voltage-dependent  $\text{Na}^{+}$  channels, prevents formation of postsynaptic GlyR clusters (data not shown). To determine whether activation of GlyR could be responsible for this activity dependence, we added the competitive antagonist strychnine to the culture medium: this did not affect the formation of neurites and nerve terminals enriched in synaptic vesicle antigens, but the number of GlyR clusters was markedly reduced and immunoreactivity of antibody 4a was evident in large intracellular structures that could be endosomes (Fig. 1a). Likewise, the GlyR-associated peripheral membrane protein gephyrin revealed by monoclonal antibody 7a (ref. 15) was no longer detected at postsynaptic sites after strychnine treatment but was seen instead at sites of focal cell attachment (Fig. 1b), as described for the early stages in development<sup>9</sup>. The overall content of GlyR and gephyrin was not reduced by addition of strychnine, as shown by dot immunoassay (Fig. 1c, d). Inhibition of GlyR clustering by strychnine was reversible after washout of strychnine at day 8 *in vitro*, and if strychnine was added after 8 days *in vitro*, when GlyR clusters had already formed, the postsynaptic localization of antigens detected

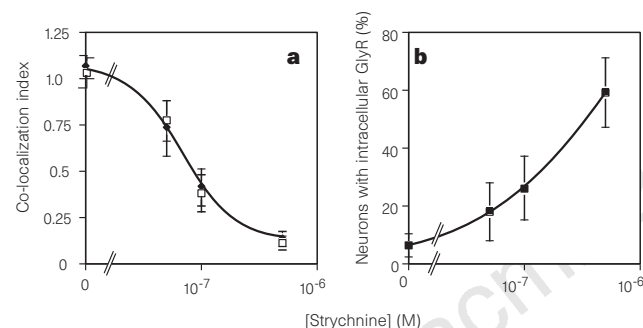


**Figure 1** Effect of strychnine treatment on the distribution of GlyR and gephyrin in spinal cultures.

**a, b.** Confocal optical sections of embryonic spinal cord neurons cultured for 8 days in the presence of 0.5  $\mu\text{M}$  strychnine. Cells were doubly immunolabelled with anti-SVP (green) and monoclonal antibodies (mAbs) 4a (red; **a**) or 7a (red; **b**). Upon strychnine treatment, mAb 4a immunoreactivity was confined to large intracellular structures (**a**, arrowheads), whereas mAb 7a immunoreactivity was concentrated at intracellular sites close to the substrate (**b**). For controls, see ref. 14. Scale bar: 5  $\mu\text{m}$ . **c, d.** GlyR and gephyrin protein levels in strychnine-treated and control cultures. The amounts of GlyR (**c**) and gephyrin (**d**), determined by mAb 4a or 7a dot immunoassay, in cultures treated with 0.05  $\mu\text{M}$  (light grey bars), 0.1  $\mu\text{M}$  (white bars) or 0.5  $\mu\text{M}$  (dark grey bars) strychnine were not significantly different from those in control cultures (black bars).

by antibodies 7a and 4a remained the same for one week after treatment (data not shown). We conclude that activation of GlyR is required for receptor clustering but not for the maintenance of differentiated postsynaptic sites.

Quantitative analysis of the fraction of sites immunoreactive against antibody 4a that were apposed to presynaptic terminals, defined here as a co-localization index, revealed that under control conditions each presynaptic nerve terminal was apposed to a GlyR cluster after 8 days *in vitro*; the co-localization index ( $\pm$  s.d.) determined from three independent experiments was  $1.03 \pm 0.08$  (Fig. 2a). Strychnine treatment drastically reduced co-localization index (half-maximum reduction at  $\sim 60$  nM strychnine (Fig. 2a), a value comparable to that which produces half-maximal inhibition of glycine responses in cultured neurons<sup>17</sup>). In contrast, the fraction of neurons with strong intracellular GlyR immunoreactivity (Fig. 1a) increased with strychnine dosage from  $7 \pm 4\%$  (control) to  $59 \pm 12\%$  at  $0.5 \mu\text{M}$  strychnine (Fig. 2b). Similar vesicles immunoreactive with 4a were also observed (Fig. 3a), when GlyR clustering was prevented by L-type  $\text{Ca}^{2+}$ -channel blockers (see later). This suggests that glycine receptors not used for synapse formation are removed by an endocytotic mechanism. This mechanism appears to

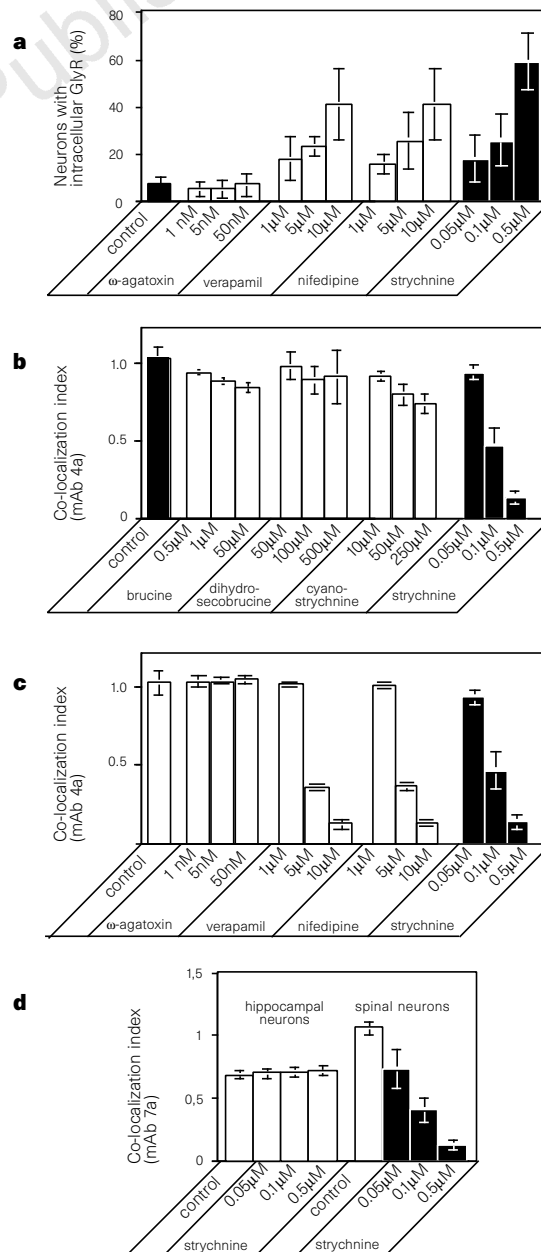


**Figure 2** Dose-response curves showing the effect of strychnine on the synaptic localization and number of cells with intracellular accumulation of GlyR immunoreactivity. **a**, Strychnine causes a dose-dependent decrease in the co-localization indices of both mAb 4a (squares) and mAb 7a (diamonds) immunoreactivity. Note the parallel reduction for both antigens. **b**, Fraction of cells showing intracellular mAb 4a immunoreactivity increases with increasing strychnine concentration.

**Figure 3** Effects of  $\text{Ca}^{2+}$ -channel blockers and strychnine analogues on GlyR internalization and synaptic targeting. **a**, Effect of  $\text{Ca}^{2+}$ -channel blockers on the number of cells with intracellular GlyR immunoreactivity. Verapamil or nifedipine increased the number of cells containing internalized GlyR to an extent similar to strychnine, whereas the P-type channel blocker  $\omega$ -agatoxin had little effect. **b**, The strychnine analogue brucine only moderately decreased the mAb 4a co-localization index at concentrations up to  $50 \mu\text{M}$ ; dihydro-seco-brucine and cyanostrychnine were even less effective. **c**, Treatment with  $\text{Ca}^{2+}$ -channel blockers inhibits GlyR-cluster formation.  $\omega$ -Agatoxin did not alter the mAb 4a co-localization index, whereas  $10 \mu\text{M}$  verapamil and nifedipine reduced it to values comparable to those obtained by treatment with  $0.5 \mu\text{M}$  strychnine. **d**, Formation of gephyrin clusters in hippocampal neurons is not altered by strychnine treatment, whereas spinal neurons show a dose-dependent decrease in gephyrin targeting. The different co-localization indices in hippocampal and spinal control cultures probably reflect the higher incidence of excitatory inputs on hippocampal neurons.

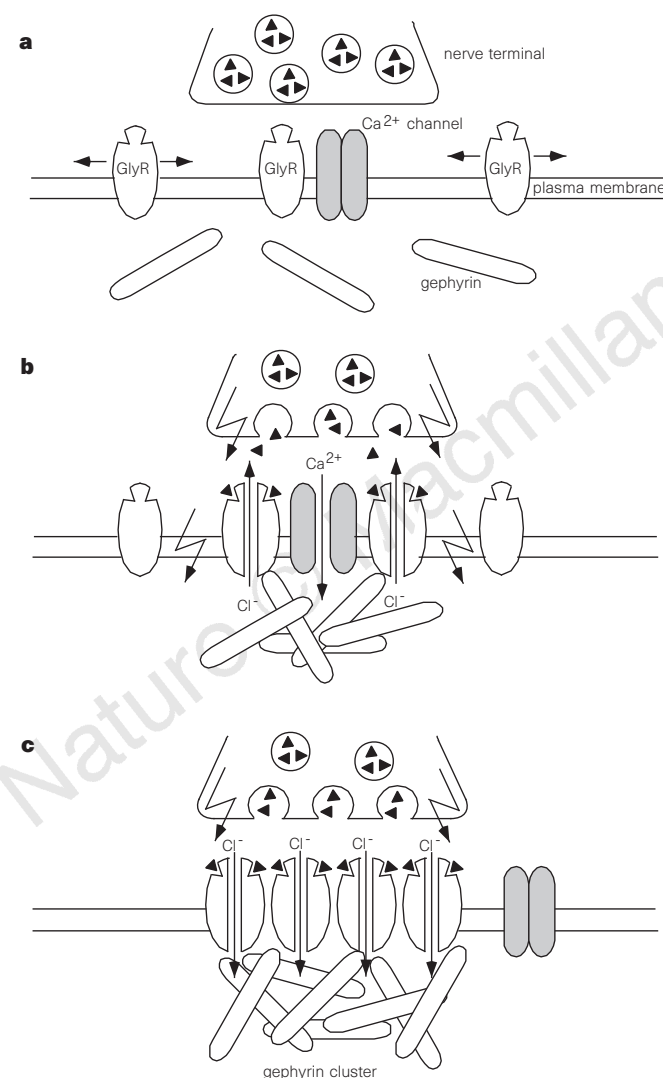
be selective, because the synaptic localization of GABA<sub>A</sub> receptors, as revealed by immunostaining with monoclonal antibody bd17 (ref. 18), was unaffected by strychnine treatment (data not shown). Also, several strychnine analogues that are moderately or slightly antagonistic to GlyR (brucine, 21,22-dihydro-23,24-seco-brucine and 19-cyano-21,22-dihydrostrychnine<sup>19,20</sup>), only decreased the 4a co-localization index to a small extent (2–6%) even at high concentrations ( $50$ – $500 \mu\text{M}$ ) (Fig. 3b). We conclude that the loss of synaptic GlyR and the intracellular accumulation of GlyR antigen both result from the antagonistic effect of strychnine on GlyR activation.

How could activation of GlyR generate a signal for gephyrin clustering in immature neurons but not after synaptic differentiation? Electrophysiology indicates that, owing to an altered  $\text{Cl}^-$  equilibrium potential, glycine responses in embryonic neurons are excitatory rather than inhibitory and can trigger  $\text{Ca}^{2+}$  influx through voltage-gated  $\text{Ca}^{2+}$  channels<sup>21–23</sup>. To test whether  $\text{Ca}^{2+}$  influx is needed for GlyR clustering, we treated spinal cord cultures with  $\text{Ca}^{2+}$ -channel antagonists that have little or no effect on presynaptic N-type  $\text{Ca}^{2+}$  channels, namely  $\omega$ -agatoxin, verapamil and nifedipine. At the concentrations used, these drugs did not



affect the morphological differentiation of cultured spinal neurons (data not shown). Treatment with the P-type  $\text{Ca}^{2+}$ -channel antagonist  $\omega$ -agatoxin did not alter the co-localization index of antibody 4a immunoreactivity or the number of cells with intercellular GlyR; however, both verapamil and nifedipine, which preferentially block L-type  $\text{Ca}^{2+}$  channels, changed these parameters in a dose-dependent fashion (Fig. 3a, c). These results indicate that  $\text{Ca}^{2+}$  influx through L-type  $\text{Ca}^{2+}$  channels is essential for GlyR clustering in cultured spinal neurons and confirm that reduced postsynaptic accumulation correlates with increased intracellular GlyR immunoactivity.

As gephyrin appears at developing postsynaptic sites before the GlyR<sup>9,12</sup>, we investigated whether the postsynaptic accumulation of gephyrin could also be affected by strychnine and  $\text{Ca}^{2+}$ -channel blockers. We found that under all conditions tested, the co-localization indices determined with antibody 7a were almost identical to those from antibody 4a, as evidenced by an index ratio of  $\sim 1$  (Table 1).



**Figure 4** A model for activity-dependent GlyR clustering at developing postsynaptic sites. **a**, Immature stage: GlyRs freely diffusing in the plasma membrane randomly distribute under a contacting presynaptic terminal. **b**, Differentiation stage: glycine release from the presynaptic terminal gates adjacent GlyRs. The resulting  $\text{Cl}^-$  efflux depolarizes the membrane and activates voltage-dependent  $\text{Ca}^{2+}$  channels, generating microdomains of increased  $\text{Ca}^{2+}$  concentration in the subsynaptic cytosol. This allows gephyrin aggregation beneath the presumptive postsynaptic membrane. **c**, Mature stage: gephyrin aggregation leads to an accumulation of GlyRs underneath the presynaptic terminal. Changes in intracellular  $\text{Cl}^-$  concentration have inverted the direction of  $\text{Cl}^-$  flux from excitatory to inhibitory. For further explanation, see text.

Thus, both antigens always co-localize at synaptic sites of spinal neurons. By contrast, in cultured hippocampal neurons, which form numerous postsynaptic gephyrin clusters but lack synaptic GlyRs<sup>24</sup>, strychnine treatment did not interfere with the formation of gephyrin clusters (Fig. 3d). The co-localization indices for 7a in these cultures were consistently  $\sim 0.7$  at all strychnine concentrations tested, indicating that gephyrin accumulation at glycinergic but not GABAergic<sup>24</sup> synapses is prevented by strychnine.

Our data support a model for the local accumulation of gephyrin and GlyR under glycinergic nerve terminals that assigns transmitter-induced activity a pivotal role in receptor clustering (Fig. 4). At early stages of development, newly synthesized GlyRs are randomly incorporated into the somatodendritic plasma membrane<sup>9,11</sup> (Fig. 4a). Activation of these receptors by glycine released from contacting presynaptic terminals causes membrane depolarization<sup>21,22</sup> and local opening of L-type  $\text{Ca}^{2+}$  channels, followed by accumulation of gephyrin at sites of  $\text{Ca}^{2+}$  influx (Fig. 4b). This localization process is likely to require additional factors released from the presynaptic terminal, because addition of glycine (at 250  $\mu\text{M}$ ) to tetrodotoxin-treated cultures fails to induce gephyrin and GlyR clusters (results not shown). The newly formed gephyrin clusters<sup>9,14</sup> are proposed to 'trap' GlyRs that laterally diffuse in the plane of the membrane and thus concentrate receptors underneath active presynaptic terminals by interaction with the GlyR  $\beta$ -subunit<sup>25,26</sup> (Fig. 4c). A decrease in intracellular  $\text{Cl}^-$  concentration then converts glycine-gated responses from excitatory to inhibitory<sup>21-23</sup>. The GlyR aggregates may subsequently be stabilized by other mechanisms, including anchoring by gephyrin to cytoskeletal elements<sup>7,14</sup>; receptors not bound by gephyrin could be removed by endocytosis, as observed in strychnine-treated spinal neurons. Our model differs significantly from that proposed for postsynaptic membrane formation at the neuromuscular junction, where presynaptically released molecules like agrin induce nicotinic acetylcholine-receptor clustering independently of activity<sup>27</sup>. This difference may arise because neurons have to establish multiple and functionally different synapses on their cell surface and therefore need stringent proof-reading mechanisms to select between functional and non-functional postsynaptic membrane specializations<sup>28</sup>. □

## Methods

**Cell culture.** Primary cultures of rat embryonic (E14) spinal cord were prepared as described<sup>9,14</sup>. Between 12 and 24 h after plating, alkaloids were added to the culture medium to the final concentrations indicated. Primary hippocampal cultures from newborn rats were prepared as described<sup>29</sup> and plated at a cell density of 80,000 cells per well. Strychnine and brucine were purchased from Merck; 21,22-dihydro-23,24-socobrucine and 19-cyano-21,22-dihydrostrychnine were donated by Hoffmann-LaRoche; the  $\text{Ca}^{2+}$ -channel blockers  $\omega$ -agatoxin, verapamil and nifedipine were from Calbiochem. Cell growth and morphological differentiation were monitored daily using phase-contrast microscopy. The total content of mAbs 4a and 7a in cultures was determined by dot immunoassay<sup>9</sup>.

**Immunocytochemistry.** Spinal cord neurons and hippocampal cultures were

**Table 1** Effect of GlyR and  $\text{Ca}^{2+}$ -channel antagonists on formation of gephyrin clusters

| Antagonist                               | Co-localization index (mAb 7a) | Co-localization index ratio* |
|------------------------------------------|--------------------------------|------------------------------|
| None                                     | 1.06 $\pm$ 0.05                | 0.97 $\pm$ 0.04              |
| Strychnine (0.5 $\mu\text{M}$ )          | 0.13 $\pm$ 0.04                | 0.89 $\pm$ 0.21              |
| Brucine (50 $\mu\text{M}$ )              | 0.83 $\pm$ 0.03                | 1.01 $\pm$ 0.02              |
| Dihydro-socobrucine (500 $\mu\text{M}$ ) | 0.84 $\pm$ 0.09                | 0.97 $\pm$ 0.01              |
| Cyanostrychnine (250 $\mu\text{M}$ )     | 0.75 $\pm$ 0.05                | 0.94 $\pm$ 0.06              |
| $\omega$ -Agatoxin (50 nM)               | 1.05 $\pm$ 0.02                | 1.01 $\pm$ 0.05              |
| Verapamil (10 $\mu\text{M}$ )            | 0.14 $\pm$ 0.02                | 0.93 $\pm$ 0.11              |
| Nifedipine (10 $\mu\text{M}$ )           | 0.13 $\pm$ 0.03                | 1.02 $\pm$ 0.03              |

\* Co-localization indices were determined on parallel cultures with both mAb 4a (not shown) and mAb 7a and were used to calculate the co-localization index ratios. A value of 1.00 for the mAb 4a/mAb 7a ratio indicates precise co-localization of both immunoreactivities. The data represent the mean  $\pm$  s.d. of three independent experiments.



processed for mAb 7a or mAb 4a (ref. 15) and anti-SVP<sup>16</sup> double immunofluorescence after 8 or 18 days in culture, respectively<sup>14</sup>. Optical sections of 0.1 µm pixel size were acquired with a Sarastro 2000 confocal laser scanning microscope (Molecular Dynamics).

**Data analysis.** For determining the co-localization index, 10 visual fields were randomly selected for each experiment. In each field the numbers of postsynaptic membrane specializations decorated with mAb 7a or mAb 4a, respectively, and of anti-SVP-positive nerve terminals were counted; their ratio was defined as the co-localization index. The co-localization index ratio was calculated by dividing the co-localization index determined with mAb 4a by that obtained with mAb 7a. Similarly, the relative number of neurons with intracellular mAb 4a immunoreactivity was counted in 10 randomly selected visual fields. All data shown represent the mean  $\pm$  s.d. of 3 independent experiments; about 100 neurons were analysed in each experiment.

Received 21 November 1997; accepted 29 January 1998.

1. Froehner, S. C. Regulation of ion channel distribution at synapses. *Annu. Rev. Neurosci.* **16**, 347–368 (1993).
2. Ehlers, M. D., Mammen, A. L., Lau, L.-F. & Hagan, R. L. Synaptic targeting of glutamate receptors. *Curr. Opin. Cell Biol.* **8**, 484–489 (1996).
3. Kirsch, J., Meyer, G. & Betz, H. Synaptic targeting of ionotropic neurotransmitter receptors. *Mol. Cell. Neurosci.* **8**, 93–98 (1996).
4. Sheng, M. Glutamate receptors put in their place. *Nature* **386**, 221–223 (1997).
5. Triller, A., Cluzeaud, F., Pfeiffer, E., Betz, H. & Korn, H. Distribution of glycine receptors at central synapses: an immunoelectron microscopy study. *J. Cell Biol.* **101**, 683–688 (1985).
6. Altschuler, R. A., Betz, H., Parakkal, M., Reeks, K. & Wenthold, R. Identification of glycinergic synapses in the cochlear nucleus through immunocytochemical localization of the postsynaptic receptor. *Brain Res.* **369**, 316–320 (1986).
7. Kirsch, J. et al. The 93-kDa glycine receptor-associated protein binds to tubulin. *J. Biol. Chem.* **266**, 22242–22245 (1991).
8. Prior, P. et al. Primary structure and alternative splice variants of gephyrin, a putative glycine receptor-tubulin linker protein. *Neuron* **8**, 1161–1170 (1992).
9. Kirsch, J., Wolters, I., Triller, A. & Betz, H. Gephyrin antisense oligonucleotides prevent glycine receptor clustering in spinal neurons. *Nature* **366**, 745–748 (1993).
10. Kuhse, J., Betz, H. & Kirsch, J. The inhibitory glycine receptor: architecture, synaptic localization and molecular pathology of a postsynaptic ion channel complex. *Curr. Opin. Neurobiol.* **5**, 318–323 (1995).
11. Hoch, W., Betz, H. & Becker, C.-M. Primary cultures of mouse spinal cord express the neonatal isoform of the inhibitory glycine receptor. *Neuron* **3**, 339–348 (1990).
12. Béchade, C., Colin, I., Kirsch, J., Betz, H. & Triller, A. Glycine receptor  $\alpha$  subunit and gephyrin expression in cultured spinal neurons: a quantitative analysis. *Eur. J. Neurosci.* **8**, 429–435 (1996).
13. Bormann, J., Hamill, O. P. & Sakmann, B. Mechanism of anion permeation through channels gated by glycine and  $\gamma$ -aminobutyric acid in mouse cultured spinal neurones. *J. Physiol. (London)* **385**, 243–286 (1987).
14. Kirsch, J. & Betz, H. The postsynaptic localization of the glycine receptor-associated protein gephyrin is regulated by the cytoskeleton. *J. Neurosci.* **15**, 4148–4156 (1995).
15. Pfeiffer, E., Simler, R., Grenningloh, G. & Betz, H. Monoclonal antibodies and peptide mapping reveal structural similarities between the subunits of the glycine receptor of rat spinal cord. *Proc. Natl Acad. Sci. USA* **81**, 7224–7227 (1984).
16. Knaus, P., Betz, H. & Rehm, H. Expression of synaptophysin during postnatal development of the mouse brain. *J. Neurochem.* **47**, 1302–1304 (1986).
17. Virginio, C. & Cherubini, E. Glycine-activated whole cell and single channel currents in rat cerebellar granule cells in culture. *Brain Res.* **98**, 30–40 (1997).
18. Fritschy, J. M. et al. Five subtypes of type A  $\gamma$ -aminobutyric acid receptors identified in neurons by double and triple immunofluorescence staining with subunit specific antibodies. *Proc. Natl Acad. Sci. USA* **89**, 6726–6730 (1992).
19. Betz, H. The glycine receptor of rat spinal cord: exploring the site of action of the plant alkaloid strychnine. *Angew. Chem. Int. Ed.* **24**, 365–370 (1985).
20. Becker, C.-M. in *Handbook of Experimental Pharmacology*, Vol. 102, *Selective Neurotoxicity* (eds Herken, H. & Hucho, F.) Vol. 1, 540–575 (Springer, Heidelberg, 1992).
21. Reichling, D. B., Kyrozis, A., Wang, J. & MacDermott, A. B. Mechanisms of GABA and glycine depolarization-induced calcium transients in rat dorsal horn neurons. *J. Physiol. (London)* **476**, 411–421 (1994).
22. Wang, J., Reichling, D. B., Kyrozis, A. & MacDermott, A. B. Developmental loss of GABA- and glycine-induced depolarization and  $Ca^{2+}$  transients in embryonic rat dorsal horn neurons in culture. *Eur. J. Neurosci.* **6**, 1275–1280 (1994).
23. Boehm, S., Harvey, R. J., v.Holst, A., Rohrer, H. & Betz, H. Glycine receptors in cultured chick sympathetic neurons are excitatory and trigger neurotransmitter release. *J. Physiol. (London)* **504**, 683–694 (1997).
24. Craig, A. M., Banker, G., Chang, W., McGrath, M. E. & Serpinsky, A. S. Clustering of gephyrin at GABAergic but not glutamatergic synapses in cultured rat hippocampal neurons. *J. Neurosci.* **16**, 3166–3177 (1996).
25. Kirsch, J., Kuhse, J. & Betz, H. Targeting of glycine receptor subunits to gephyrin-rich domains in transfected human embryonic kidney cells. *Mol. Cell. Neurosci.* **6**, 450–461 (1995).
26. Meyer, G., Kirsch, J., Betz, H. & Langosch, D. Identification of a gephyrin binding motif on the glycine receptor  $\beta$  subunit. *Neuron* **15**, 563–572 (1995).
27. Hall, Z. W. & Sanes, J. R. Synaptic structure and development: the neuromuscular junction. *Cell* **72** (suppl.), 99–121 (1993).
28. Changeux, J. P. & Danchin, A. Selective stabilization of developing synapses as a mechanism for the specification of neuronal networks. *Nature* **264**, 705–712 (1976).
29. Goslin, K. & Banker, G. in *Culturing Nerve Cells* (eds Banker, G. & Goslin, K.) 251–282 (MIT Press, Cambridge, MA, 1991).

**Acknowledgements.** We thank I. Bartnik for technical assistance and V. O'Connor, R. Harvey and A. Puschel for critically reading the manuscript. This work was supported by grants from the Deutsche Forschungsgemeinschaft and Fonds der Chemischen Industrie. J.K. holds an endowed professorship from the Stifterverband für die Deutsche Wissenschaft.

Correspondence and requests for materials should be addressed to J.K. (e-mail: neurochemie@mpih-frankfurt.mpg.de).

## Cryptochrome blue-light photoreceptors of *Arabidopsis* implicated in phototropism

Margaret Ahmad, Jose A. Jarillo, Olga Smirnova & Anthony R. Cashmore

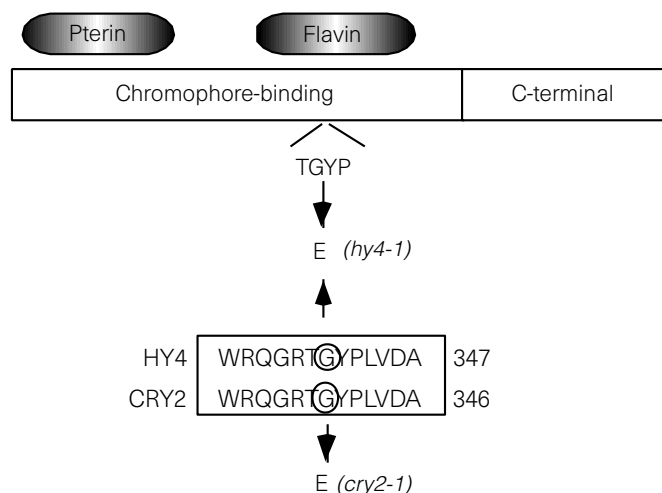
Plant Science Institute, Department of Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6018, USA

Phototropism—bending towards the light—is one of the best known plant tropic responses<sup>1,2</sup>. Despite being reported by Darwin and others<sup>3,4</sup> over a century ago to be specifically under the control of blue light, the photoreceptors mediating phototropism have remained unknown. We have characterized a blue-light photoreceptor from *Arabidopsis*, named CRY1 for cryptochrome 1 (ref. 5); this photoreceptor is a flavoprotein that mediates numerous blue-light-dependent responses<sup>6</sup>. In *Arabidopsis*, *HY4* (the gene encoding CRY1) is a member of a small gene family that also encodes a related photoreceptor, CRY2 (refs 7, 8), which shares considerable functional overlap with CRY1 (ref. 9). Here we report that mutant plants lacking both the CRY1 and the CRY2 blue-light photoreceptors are deficient in the phototropic response. Transgenic *Arabidopsis* plants overexpressing CRY1 or CRY2 show enhanced phototropic curvature. We conclude that cryptochrome is one of the photoreceptors mediating phototropism in plants.

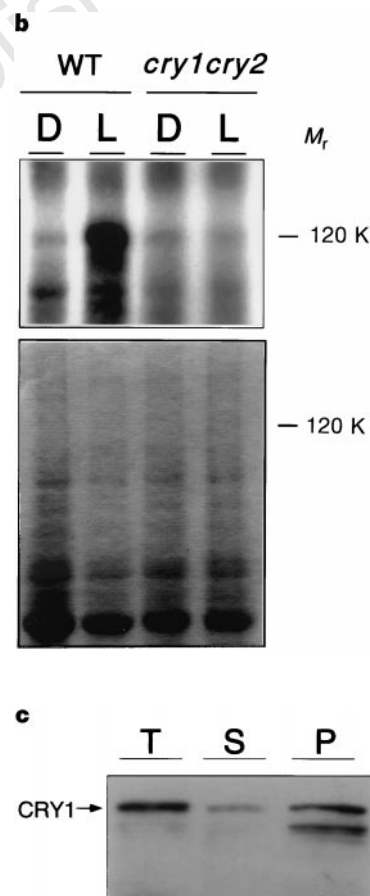
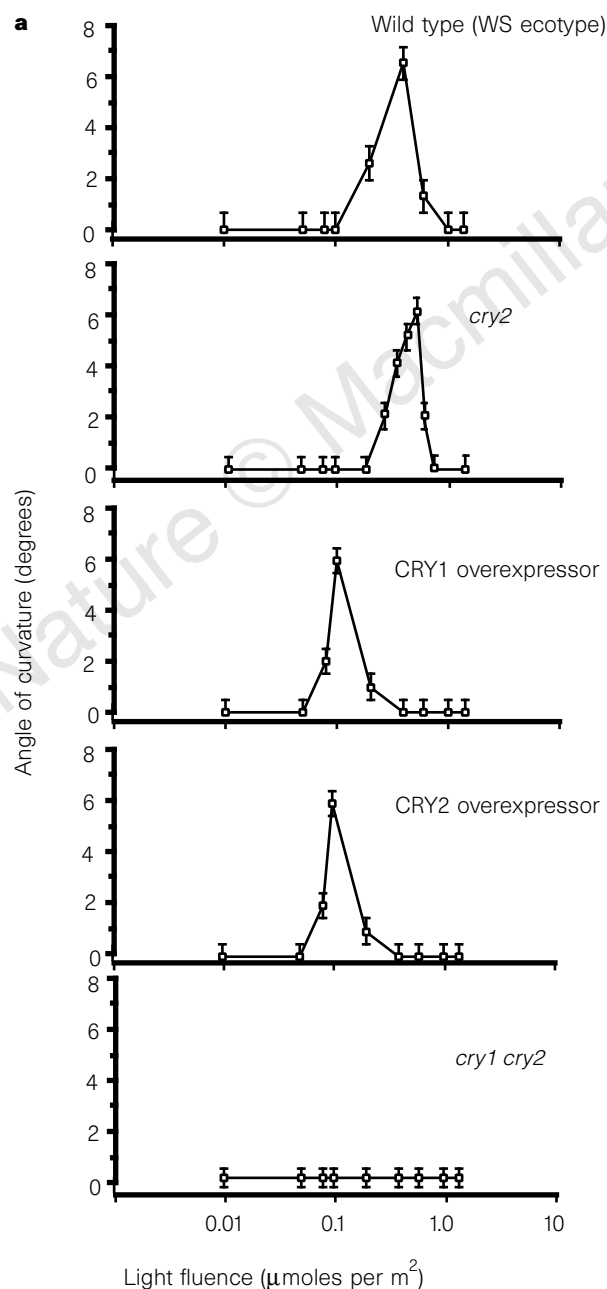
Many features of CRY1-mediated responses show similarity to features associated with phototropism, including similar action spectra<sup>10,11</sup>, predicted multiple chromophores (including a flavin)<sup>12–14</sup>, and enhancement of activity by phytochrome<sup>15–17</sup>. Although CRY1-deficient (*hy4* mutant)<sup>18</sup> plants can show phototropic responses, we reasoned that, if cryptochrome is the receptor mediating phototropism, both CRY1 and CRY2 might be able to mediate this response, and only by eliminating both homologues would considerable impairment of phototropism be seen. Accordingly, we screened for second-site mutations in the *hy4* (CRY1-deficient) genetic background to identify possible mutations in CRY2 that lead to loss of the phototropic response. Ten aphototropic mutant lines were isolated by screening 20,000 M2 seeds that had been treated with mutagen, and the CRY2 gene in each line was sequenced.

In one line we identified a point mutation resulting in an amino-acid substitution (Gly 340  $\rightarrow$  Glu) in a highly conserved motif of the flavin-binding domain (Fig. 1). A mutation in the identical residue of CRY1 (producing the *hy4-1* mutant) had been previously identified and shown to fully inactivate this receptor<sup>5,19</sup>. We designate this new mutant allele *cry2-1*. To confirm that the absence of first-positive phototropism was correlated with mutations in both CRY1 and CRY2, the double mutant was crossed to wild-type plants and selfed, and 312 F<sub>2</sub> progeny were then scored for phototropism. We identified 17 seedlings that did not show phototropism, consistent with segregation of the non-phototropic phenotype with the double (*cry1cry2*) mutation ( $\chi^2 = 0.013$ ). Non-phototropic seedlings were further examined for the presence of CRY1 (by western blotting) and CRY2 (by sequencing the CRY2 locus). All 17 seedlings lacked CRY1 protein and contained the mutant allele of CRY2.

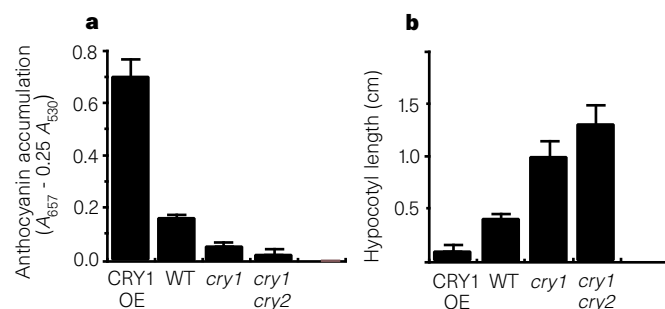
We analysed the effect of cryptochrome on phototropism in cryptochrome-overexpressing and null-mutant seedlings (Fig. 2a). The *cry1cry2* double mutant showed no evidence of first-positive curvature. In contrast, transgenic *Arabidopsis* seedlings overexpressing cryptochrome genes showed hypersensitivity to blue light during first-positive phototropism, with a peak of activity



**Figure 1** Identification of a mutation in the gene encoding CRY2. The domains of the protein are shown at the top. Flavin and pterin are chromophores. Genomic DNA from a non-phototropic mutant line was amplified by the polymerase chain reaction and sequenced. A single base-pair substitution (G → A) was identified in the CRY2-coding sequence. This substitution results a change from Gly34 to Glu (this mutant is the *cry2-1* allele) in a conserved region of the flavin-binding domain. Mutation of the precise homologous sequence was previously identified in the CRY1-coding sequence of the *hy4-1* allele<sup>5</sup>. Part of the sequences of CRY1 (HY4) and CRY2 are boxed near the bottom of the figure. 347 and 346 indicate the position of the last amino acids of these sequences in the entire amino-acid sequence of these proteins.



**Figure 2** Effect of cryptochrome on the phototropic response. **a**, First-positive curvature in cryptochrome-overexpressing and -deficient mutant seedlings. Three independently treated plates were used per data point; the average curvature from each plate was calculated; error bars represent the mean of the different between the average value from the three plates. **b**, *In vitro* blue-light-dependent phosphorylation of the 120K protein in membrane fractions prepared from wild-type (WS ecotype) and *cry1cry2* mutant seedlings. The top panel represents the autoradiograph; the bottom panel represents the dried, Coomassie-stained protein gel. D, mock-irradiated control; L, blue-light-irradiated plants. **c**, Western blot analysis of levels of CRY1 proteins in fractions prepared for *in vitro* phosphorylation assays. Portions were removed from the following fractions: T, total; S, supernatant; and P, pellet, as described in the Methods.



**Figure 3** Anthocyanin accumulation and inhibition of hypocotyl elongation (blue-light responses) in *cry* mutant and CRY1-overexpressing seedlings. **a**, Anthocyanin accumulation in mutant and CRY1-overexpressing seedlings (CRY1 OE). WT, wild-type. **b**, Inhibition of hypocotyl elongation under blue light in mutant and CRY1-overexpressing seedlings.

(~0.1  $\mu$ mol of fluence) occurring at fivefold lower light fluence than in the wild type (~0.5  $\mu$ mol). This enhancement concurs with the degree of enhanced responsivity to blue light seen for anthocyanin accumulation and inhibition of hypocotyl elongation in cryptochrome-overexpressing lines<sup>20</sup>, and indicates that cryptochrome may also function as the receptor for phototropism. Mutant seedlings lacking CRY1 showed little difference to wild-type seedlings in first-positive curvature<sup>21</sup> (not shown). Interestingly, *cry2* mutant seedlings also showed little impairment of first-positive curvature (Fig. 2a), indicating that wild-type levels of either CRY1 or CRY2 alone are enough to mediate a robust phototropic response. This situation is reminiscent of that found with a number of low-fluence responses mediated by phytochrome, such as induction of *CAB* gene expression<sup>22</sup>. Gene expression is normal in either *phyA* or *phyB* single mutants, but dramatically reduced in the *phyAphyB* double mutant, indicating that levels of phytochrome could be required to fall below a certain (low) threshold to dramatically affect responses requiring low light intensity.

There was little difference in the extent of second-positive curvature among the various mutant and cryptochrome-overexpressing lines. After repeated backcrossing, the *cry1cry2* double mutant also retained some second-positive curvature, indicating the presence of other blue-light-absorbing species implicated in phototropism, in addition to the cryptochromes. Given their effect on first-positive phototropism, however, we conclude that phototropic curvature is mediated by cryptochromes, and that either CRY1 or CRY2 can function as a photoreceptor.

Blue-light-dependent phosphorylation of a membrane-associated protein of relative molecular mass 120K has been linked to phototropism by identification of a phototropic-null mutant (*nph1*) that lacks this particular response (refs 23, 24). *NPH1* lies on chromosome III (ref. 25) and is therefore distinct from CRY2, which lies on chromosome I between the *CER1* and *PHYA* genetic markers and has been mapped to a bacterial artificial chromosome clone (F19P19; Genbank accession number AC000104). It has been proposed that *NPH1* might encode the blue-light photoreceptor for phototropism, and that *NPH1* might undergo blue-light-dependent autophosphorylation<sup>24,25</sup>. We investigated the role of cryptochrome in the blue-light-dependent autophosphorylation of *NPH1* protein in the *cry1cry2* mutant. A high-molecular-weight phosphorylated band of ~120K was observed in both wild-type and *cry1cry2* mutant seedlings, indicating that *NPH1* levels are not affected by lack of cryptochrome. However, light-dependent phosphorylation of this band was absent in *cry1cry2* seedlings at light intensities significant for phototropism, in contrast to the situation in wild-type *Arabidopsis* under these conditions (Fig. 2b). Western blot analysis confirmed that CRY1 is enriched in the membrane fractions used in these experiments (Fig. 2c). Therefore, phosphorylation of the 120K protein that is dependent on low-fluence blue light seems to result from the action of cryptochrome (either CRY1 or CRY2). *NPH1* encodes a protein kinase with a putative redox-sensing domain<sup>25</sup>; if it proves to be a flavoprotein it might serve as acceptor

for light-dependent electron transfer mediated by CRY1 or CRY2. Surprisingly, none of the previously described phototropic mutants, including *nph1*, significantly affects other cryptochrome-associated phenotypes such as inhibition of stem elongation<sup>21,24</sup>. Phototropism may represent a specialized branch of the blue-light-sensing pathway, with a unique transduction pathway not shared by other responses to blue light. This would help to explain the many distinct physiological characteristics of the phototropic response.

In addition to being deficient in phototropism, the *cry1cry2* mutant was more impaired in blue-light responses such as anthocyanin accumulation and inhibition of hypocotyl elongation than was the CRY1-deficient parent<sup>19,20</sup> (Fig. 3a, b). These observations correspond with the marked degree of functional overlap observed between the CRY1 and CRY2 blue-light photoreceptors, as shown by overexpression and fusion protein studies<sup>9</sup>.

Clearly many blue-light responses are mediated by cryptochrome. These include inhibition of hypocotyl elongation, accumulation of anthocyanin, leaf and cotyledon expansion, extension growth, petiole elongation, gene expression, promotion of flowering, membrane depolarization<sup>6,26</sup> and now phototropism. The cryptochrome gene family is large and evolutionarily conserved, and contains many members (CRY1, CRY2 and others) in the many different higher and lower plant species<sup>6,27</sup>. Even though additional blue-light-absorbing species contribute to blue-light responses such as phototropism, the cryptochromes seem to be the predominant blue-light photoreceptors in plants. □

## Methods

**Plant growth and genetic screening.** Routine growth and maintenance of plants and mutagenesis with ethyl-methane-sulphonate were carried out by standard methods<sup>28</sup>. Wild-type seedlings were of ecotype WS (Wassilewskaja); *hy4* mutant seeds used for mutagenesis were *hy4-3* (ref 5) (also ecotype WS); CRY1-overexpressing seedlings were as described<sup>20</sup>; and CRY2-overexpressing lines were as indicated<sup>9</sup>. For phototropic mutant screens, seeds were plated on MS agar medium and kept at 4 °C for two days, then transferred to bright white light for 36 h to induce germination. Plates were subsequently kept in the dark for 42 h, then transferred to unilateral blue light at a fluence of 0.1  $\mu$ mol per m<sup>2</sup> per second for 24 h. Seedlings that failed to show phototropism were identified, transferred to soil, and analysed further.

**Measurement of phototropic curvature (first-positive phototropism).** Approximately 60 seeds from the wild-type, mutant and overexpressing plants were planted in straight rows of 20 seedlings each on agar plates (1 mM KNO<sub>3</sub>, 0.8% agar) and kept at 4 °C for two days. Plates were subsequently transferred to high-fluence white light for 36 h to induce germination and then transferred to the dark for an additional 42 h. Seedlings were pulsed with 0.1  $\mu$ mol per m<sup>2</sup> per second unilateral blue light for appropriate durations to give the indicated total fluence (Fig. 2a; x-axis). For light treatments of less than 0.1  $\mu$ mol, pulses of 0.01  $\mu$ mol per m<sup>2</sup> per second were used. After light pulses, seedlings were returned to darkness and the angle of curvature was measured after 2 h as described<sup>24</sup>.

**Membrane preparations and *in vitro* phosphorylation assays.** Membrane-enriched fractions and *in vitro* phosphorylation assays (Fig. 2b) were performed as described<sup>24</sup>, except that 25  $\mu$ g crude membranes and 10  $\mu$ mol



total fluence light were used. For western blot experiments (Fig. 2c), portions were removed from the following fractions: total, 11,000g supernatant, representing clarified homogenate after removal of debris; supernatant, representing supernatant after a high-speed spin for 75 min at 100,000g; pellet, representing the membrane pellet fraction after the high-speed spin. Samples were boiled in SDS sample buffer and resolved on SDS polyacrylamide gels, and western blotting with anti-CRY1 antibody was carried out as described<sup>20</sup>. Each lane represents an equivalent proportion (1%) of the indicated fraction. The lower band appearing in the membrane pellet fraction appears to be a degradation product resulting from this protocol. Phytochrome was also enriched in this fraction (not shown). The contrast in sedimentation of CRY1 between this and other protocols<sup>20</sup> is striking, and is possibly due to differences in buffer composition (CRY1 is relatively insoluble, sticks more readily to other proteins, and is more readily pelleted under conditions of low ionic strength; M.A., unpublished observations).

**Anthocyanin accumulation and inhibition of hypocotyl elongation** were measured as described<sup>19</sup>.

Received 19 September 1997; accepted 6 February 1998.

- Iino, M. Phototropism: mechanisms and ecological implications. *Plant Cell Environ.* **13**, 633–650 (1990).
- Poff, L. K. *et al.* in *Arabidopsis* (eds Meyerowitz, E. S. & Somerville, C. R.) 639–664 (Cold Spring Harbor Laboratory Press, New York, 1994).
- Sachs, J. Wirkungen des farbigen Lichts auf Pflanzen. *Bot. Z.* **353**–358 (1864).
- Darwin, C. *The Power of Movement in Plants* (D. Appleton, New York, 1881).
- Ahmad, M. & Cashmore, A. R. *HY4* gene of *A. thaliana* encodes a protein with characteristics of a blue-light photoreceptor. *Nature* **366**, 162–166 (1993).
- Ahmad, M. & Cashmore, A. R. Seeing blue: the discovery of cryptochrome. *Plant Mol. Biol.* **30**, 851–861 (1996).
- Hoffman, P. D., Batschauer, A. & Hays, J. B. *PHH1*, a novel gene from *Arabidopsis thaliana* that encodes a protein similar to plant blue-light photoreceptors and microbial photolyses. *Mol. Gen. Genet.* **253**, 259–265 (1996).
- Lin, C., Ahmad, M., Chan, J. & Cashmore, A. R. *CRY2*, a second member of the *Arabidopsis* cryptochrome gene family. *Plant Physiol.* **110**, 1047 (1996).
- Ahmad, M., Jarillo, J. A. & Cashmore, A. R. Chimeric proteins between CRY1 and CRY2 *Arabidopsis* blue light photoreceptors indicate overlapping function and varying protein stability. *Plant Cell* **10**, 197–208 (1998).
- Shropshire, W. J. & Withrow, R. B. Action spectrum of phototropism tip curvature of *Avena*. *Plant Physiol.* **33**, 360–365 (1958).
- Young, J. C., Liscum, E. & Hangarter, R. P. Spectral-dependence of light-inhibited hypocotyl elongation in photomorphogenic mutants of *Arabidopsis*: evidence for a UV-A photosensor. *Planta* **188**, 106–114 (1992).
- Konjevic, R., Steinitz, B. & Poff, K. Dependence of the phototropic response of *Arabidopsis thaliana* on fluence rate and wavelength. *Proc. Natl. Acad. Sci. USA* **86**, 9876–9880 (1989).
- Malhotra, K., Kim, S. T., Batschauer, A., Dawut, L. & Sancar, A. Putative blue-light photoreceptors from *Arabidopsis thaliana* and *Sinapis alba* with a high-degree of sequence homology to DNA photolyase contain the two photolyase cofactors but lack DNA-repair activity. *Biochemistry* **34**, 6892–6899 (1995).
- Lin, C. *et al.* Association of flavin adenine dinucleotide with the *Arabidopsis* blue-light receptor CRY1. *Science* **269**, 968–970 (1995).
- Ahmad, M. & Cashmore, A. R. The blue-light receptor cryptochrome 1 shows functional dependence on phytochrome A or phytochrome B in *Arabidopsis thaliana*. *Plant J.* **11**, 421–427 (1997).
- Janoudi, A. K. *et al.* Multiple phytochromes are involved in red-light-induced enhancement of first-positive phototropism in *Arabidopsis thaliana*. *Plant Physiol.* **113**, 975–979 (1997).
- Janoudi, A. K. *et al.* Both phytochrome A and phytochrome B are required for the normal expression of phototropism in *Arabidopsis thaliana*. *Physiol. Plant.* **101**, 278–282 (1997).
- Koornneef, M., Rolff, E. & Spruit, C. J. P. Genetic control of light-inhibited hypocotyl elongation in *Arabidopsis thaliana* (L.) Heynh. *Z. Pflanzenphysiol. Bd.* **100**, 147–160 (1980).
- Ahmad, M., Lin, C. & Cashmore, A. R. Mutations throughout an *Arabidopsis* blue-light photoreceptor impair blue-light-responsive anthocyanin accumulation and inhibition of hypocotyl elongation. *Plant J.* **8**, 653–658 (1995).
- Lin, C., Ahmad, M. & Cashmore, A. R. *Arabidopsis* cryptochrome 1 is a soluble protein mediating blue-light-dependent regulation of plant growth and development. *Plant J.* **10**, 893–902 (1996).
- Liscum, E. & Briggs, W. R. Mutations of *Arabidopsis* in potential transduction and response components of the phototropic signaling pathway. *Plant Physiol.* **112**, 291–296 (1996).
- Reed, J. W., Nagatani, A., Elich, T. D., Fagan, M. & Chory, J. Phytochrome A and phytochrome B have overlapping but distinct functions in *Arabidopsis* development. *Plant Physiol.* **104**, 1139–1149 (1994).
- Khurana, J. & Poff, K. Mutants of *Arabidopsis thaliana* with altered phototropism. *Planta* **178**, 400–406 (1989).
- Liscum, E. & Briggs, W. R. Mutations in the *NPH1* locus of *Arabidopsis* disrupt the perception of phototropic stimuli. *Plant Cell* **7**, 473–485 (1995).
- Huala, E. *et al.* *Arabidopsis* NPH1: a protein kinase with a putative redox-sensing domain. *Science* **278**, 2120–2123 (1997).
- Bagnall, D. J., King, R. W. & Hangarter, R. P. Blue-light promotion of flowering is absent in *hy4* mutants of *Arabidopsis*. *Planta* **200**, 278–280 (1996).
- Small, D. B., Min, B. & Lefebvre, P. A. Characterization of a *Chlamydomonas reinhardtii* gene encoding a protein of the DNA photolyase/blue light photoreceptor family. *Plant Mol. Biol.* **28**, 443–454 (1995).
- Koncz, C., Chua, N.-H. & Schell, J. *Methods in Arabidopsis research*. (World Scientific, Singapore, 1992).

**Acknowledgements.** We thank members of our lab and of the Plant Science Institute for discussions. J.A.J. is the recipient of a Spanish Ministry of Education postdoctoral fellowship. This work was funded by a grant from the NIH to A.R.C.

Correspondence and requests for materials should be addressed to M.A. (e-mail: mahmad@mail.sas.upenn.edu).

## Control of antennal versus leg development in *Drosophila*

Fernando Casares & Richard S. Mann

Department of Biochemistry and Molecular Biophysics, Center for Neurobiology and Behavior, Columbia University College of Physicians and Surgeons, 701 West 168th Street, HHSC 1108, New York, New York 10032, USA

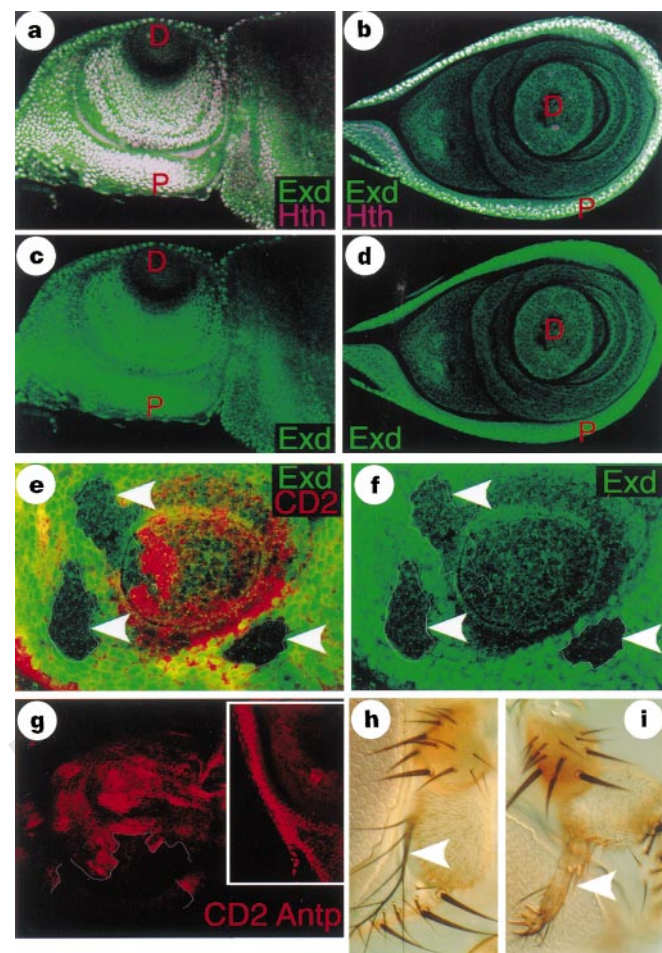
During the evolution of insects from a millipede-like ancestor, the Hox genes are thought to have promoted the diversification of originally identical body structures<sup>1,2</sup>. In *Drosophila melanogaster*, antennae and legs are homologous structures that differ from each other as a result of the Hox gene *Antennapedia* (*Antp*), which promotes leg identities by repressing unknown antennal-determining genes<sup>3–7</sup>. Here we present four lines of evidence that identify *extradenticle* (*exd*) and *homothorax* (*hth*) as antennal-determining genes. First, removing the function of *exd*<sup>8,9</sup> or *hth*, which is required for the nuclear localization of Exd protein<sup>10</sup>, transforms the antenna into leg; such transformations occur without activation of *Antp*. Second, *hth* is expressed and Exd is nuclear in most antennal cells, whereas both are restricted to proximal cells of the leg. Third, *Antp* is a repressor of *hth*. Fourth, ectopic expression of *Meis1*, a murine *hth* homologue<sup>10</sup>, can trigger antennal development elsewhere in the fly. Taken together, these data indicate that *hth* is an antennal selector gene, and that *Antp* promotes leg development by repressing *hth* and consequently nuclear Exd.

In *Drosophila*, dominant gain-of-function alleles of the Hox gene *Antp* cause the transformation of antenna into leg<sup>3,11,6,7</sup>. These transformations are due to the ectopic expression of *Antp* in the antennal imaginal disc<sup>12</sup>. Conversely, removing *Antp* from the thorax, where it is normally expressed<sup>13–15</sup>, results in the reciprocal transformation of leg into antenna<sup>4,5</sup>. These experiments led to the proposal that *Antp* represses the activity of head-determining genes, such as those that determine antenna<sup>4,5</sup>. Antennal-determining genes are predicted to be expressed in the antennal, but not in the leg, primordia (the imaginal discs); they should be required for antennal development; and they should be repressed by *Antp*. Finally, it might be possible to induce antennal development by ectopically expressing these genes. Here we show that the homeobox genes *exd* (ref. 16) and *hth* (ref. 10) fulfill all of these predictions.

In mature third-instar larvae, nuclear Exd is found in only the most proximal cells of leg imaginal discs (Fig. 1b, d), but is nuclear almost to the centre (distal cells) of antennal discs (Fig. 1a, c). The *hth* gene is expressed coincidentally with nuclear Exd in leg and antennal discs (Fig. 1a, b)<sup>10</sup>. As in embryos<sup>10</sup>, *hth* is required for the nuclear localization of Exd in imaginal discs: in clones of a strong hypomorphic allele, *hth*<sup>Cl</sup>, Exd is cytoplasmic in regions of antennal discs where it is normally nuclear (Fig. 1e, f), without affecting *exd* transcription (not shown). Conversely, ectopic expression of *hth* or *Meis1*, a murine homologue of *hth* (ref. 10), drives Exd into nuclei in cells where it is normally cytoplasmic (see below).

Removing *exd* function from a developing antenna results in a cell-autonomous transformation to leg<sup>8,9</sup>. Similarly, *hth*<sup>Cl</sup> clones in antenna are autonomously transformed to leg (Fig. 1h, i). Because ectopic *Antp* expression can also generate this transformation, we investigated whether *Antp* was de-repressed in *exd*<sup>−</sup> or *hth*<sup>−</sup> clones. Importantly, none of the three Hox genes that normally determine leg identities in *Drosophila* (*Sex combs reduced* (*Scr*), *Antp* or *Ultrabithorax* (*Ubx*)), are expressed in *exd*<sup>−</sup> or *hth*<sup>−</sup> clones in antennal discs (Fig. 1g and data not shown). Therefore, in the absence of *hth* or *exd*, leg development proceeds without Hox gene input, suggesting that the leg pathway is the 'ground state' for ventral appendage development.

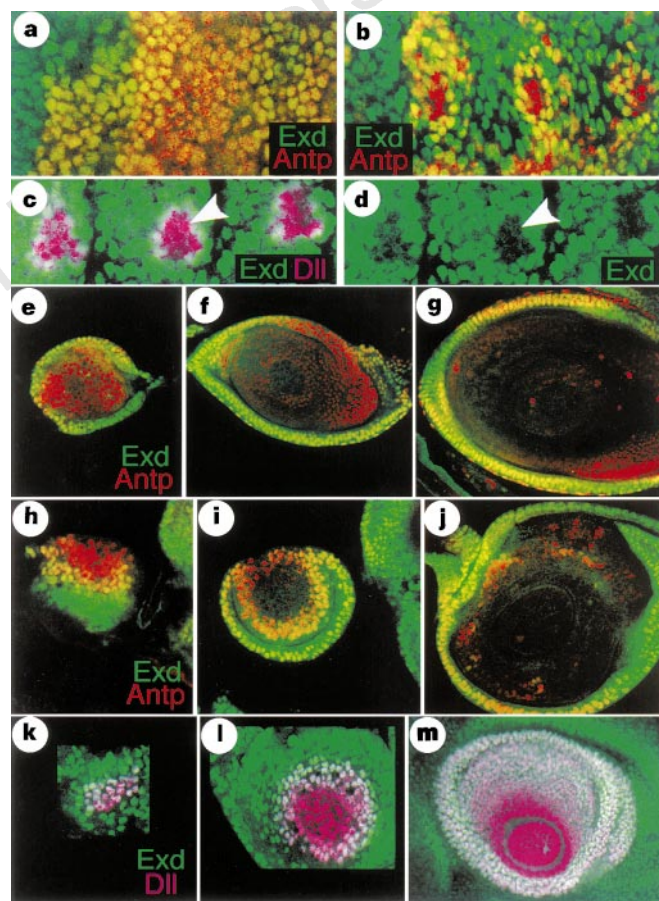
We tested whether *Antp* represses *hth* expression and the nuclear localization of Exd in three ways. First, we examined the antennal discs of *Antp*<sup>73b</sup> flies, in which ectopic *Antp* expression results in a nearly complete transformation of antenna to leg<sup>6,12</sup>. In *Antp*<sup>73b</sup> antennal discs the patterns of *Antp*, *Hth* and nuclear Exd appear indistinguishable from those observed in wild-type leg discs (Fig. 2e–j and data not shown). Second, we used the ‘flip-out’ technique<sup>17</sup> to ectopically express *Antp* in otherwise wild-type antennal discs. Wherever *Antp* is expressed in distal cells, *hth* is repressed and Exd is cytoplasmic (Fig. 3a–d). Repression of *hth* by *Antp* is cell autonomous. Third, we examined the pattern of Exd expression in leg discs containing *Antp*<sup>−</sup> clones. Before presenting these data, we describe the temporally dynamic expression pattern of *Antp* in wild-type leg discs. In mature leg discs (~120 h after egg laying), strong *Antp* expression is restricted to the most proximal cells; in the rest of the disc, *Antp* expression is weak and limited to very few cells (Fig. 2g)<sup>15</sup>.



**Figure 1** The *hth* gene is required for antennal development and the nuclear localization of Exd. **a–d**, Wild-type third-instar antenna (**a**, **c**) or leg (**b**, **d**) discs stained for Exd (green) and Hth (magenta); colocalization appears white. P and D refer to proximal and distal positions in these discs, respectively. **e**, **f**, *hth*<sup>Cl</sup> clones (arrowheads) in an antennal disc show cytoplasmic localization of Exd (green). The clones, which were detected by the absence of CD2 staining (red), are indicated by the white outlines; in **f** only Exd is shown. **g**, A *hth*<sup>Cl</sup> clone in an antennal disc stained for *Antp* (red). Clones were identified by the lack of CD2 staining (red). Although no *Antp* staining was observed in antennal discs, thoracic discs in this preparation were stained; the inset shows a T2 leg disc. Similar results were obtained with *Scr* and *Ubx* (not shown). **h**, **i**, A wild-type antenna (**h**) and an antenna with a *hth*<sup>Cl</sup> clone that is autonomously transformed to leg (**i**). The arrowheads point to an arista (**h**) or the homologous leg structure (tarsus) (**i**). *hth*<sup>Cl</sup> tissue marked by the absence of *y*<sup>+</sup>.

However, earlier in disc development (~72 h after egg laying), strong *Antp* expression is present throughout the entire leg disc; it is gradually lost from the central and middle regions of the disc as development proceeds (Fig. 2e–g). Even earlier, during embryogenesis, *Antp* is expressed in the cells that give rise to all three leg discs (Fig. 2a, b). Although Exd is nuclear in the leg disc primordia early in embryogenesis (Fig. 2a), from stage 14 onwards *hth* expression and nuclear-localized Exd are observed only in the most proximal cells of the leg primordia (Figs 1b and 2b–g). In contrast, in wild-type antennal discs, in which *Antp* is never expressed, *hth* is expressed and Exd is nuclear throughout most of the disc (Figs 1a and 2k–m).

Two characteristics of *Antp*<sup>−</sup> clones in leg discs are noteworthy. First, nuclear Exd is observed in *Antp*<sup>−</sup> clones that arise in intermediate or proximal regions of second thoracic (T2) leg discs, although in distal *Antp*<sup>−</sup> clones Exd remains cytoplasmic (Fig. 3e, f



**Figure 2** The patterns of *Antp* and nuclear Exd during development. **a**, **b**, The leg primordia regions of embryos at stage 9 (**a**) and 14 (**b**) stained for Exd (green) and *Antp* (red); colocalization appears yellow. **c**, **d**, The leg primordia region of a *Dll-lacZ/+* stage 14 embryo stained for Exd (green) and  $\beta$ -gal (*Dll*; magenta); colocalization appears white. In **d** only Exd is shown. *Dll* expression marks the leg primordia<sup>29</sup> and Exd is already cytoplasmic in the most distal cells at this early stage of leg development. **e–j** Three stages of a wild-type leg disc (**e–g**) and an *Antp*<sup>73b</sup> antennal disc (**h–j**) stained for *Antp* (red) and Exd (green). **k–m**, Three stages of a wild-type antennal disc stained for Exd (green) and *Dll* (magenta); colocalization appears white. **e**, **h**, **k** ~72 h, **f**, **i**, **l** ~96 h, and **g**, **j**, **m** ~120 h after egg laying. Because there are no clonal restrictions along the proximodistal axis in leg discs<sup>30</sup>, descendants of *Antp*<sup>−</sup>*hth*<sup>+</sup> proximal cells can contribute to both proximal and distal regions of the leg.



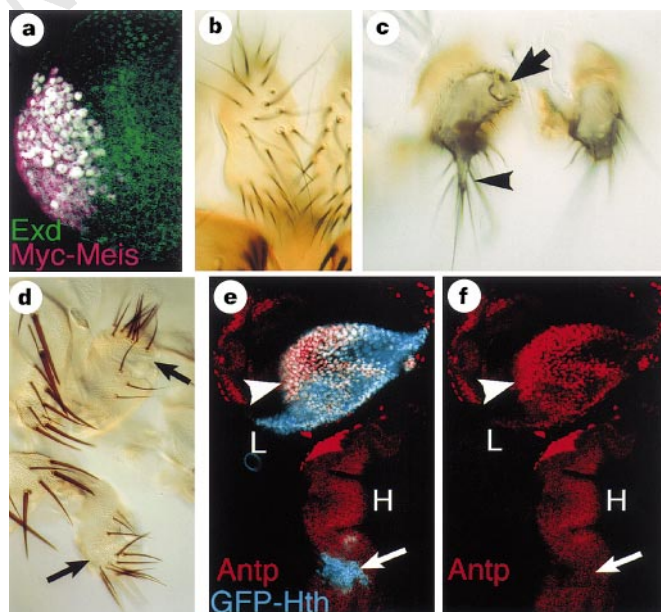
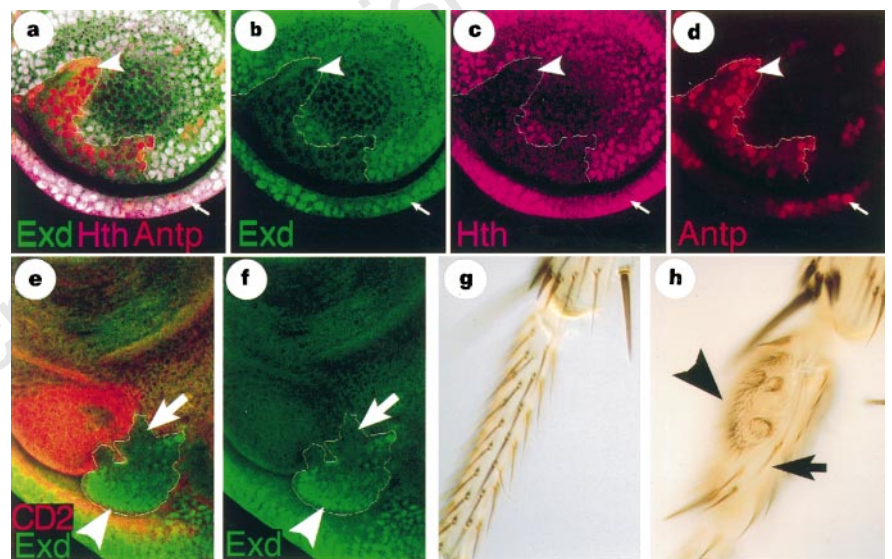
and data not shown). This difference is predicted from the dynamic nature of *Antp* expression during development: at the time when these *Antp*<sup>-</sup> clones were generated (first instar), Exd was already in the cytoplasm in the most distal cells of the disc (Fig. 2c, d) whereas in more proximal cells Exd was nuclear (Fig. 2e–g). A second characteristic of *Antp*<sup>-</sup> clones is that, even in proximal ones, only a subset of the *Antp*<sup>-</sup> cells have nuclear Exd (Fig. 3e, f). Although the reason for this heterogeneity is not clear, it is reminiscent of previous results<sup>4,5</sup> showing that only a subset of the cells in an *Antp*<sup>-</sup> clone generate antennal structures; some *Antp*<sup>-</sup> cells still form leg (Fig. 3h). This heterogeneity is unlikely to be due to a diffusible factor originating from *Antp*<sup>+</sup> cells because when *Antp* was ectopically expressed in the flip-out experiment (Fig. 3a–d) the effects on Exd localization and *hth* expression were cell autonomous.

To confirm that *hth* is an antennal-determining gene we investigated whether ectopic *hth* expression could lead to antennal development elsewhere in the fly. For this experiment, we used the Gal4 system<sup>18</sup> to ectopically express the murine homologue of *hth*, *Meis1*, in the *Distal-less* (*Dll*) domain<sup>19</sup>. *Meis1* expression in the *Dll* domain

of the anal primordium of the genital disc, where normally *hth* is not expressed and Exd is cytoplasmic, resulted in nuclear Exd and the transformation of anal plates into antennae (Fig. 4a–c). In contrast, expressing a constitutively nuclear form of Exd (by fusing it to a heterologous nuclear localization sequence), did not generate this transformation, suggesting that both *hth* and *exd* are required (data not shown). In leg discs, *Meis1* or *hth* expression in the *Dll* domain resulted in leg truncations that are similar to those generated by high levels of *exd* (ref. 20) (Fig. 4d). As well as inducing the nuclear localization of Exd, expression of *hth* or *Meis1* in leg discs resulted in the maintenance of strong *Antp* expression (Fig. 4e, f; compare with Fig. 2g). We suggest that leg truncations result because the coexpression of *Antp* and *hth* is incompatible with limb development. In contrast, like the antenna, the anal primordium normally develops without contributions from the bithorax or Antennapedia complex genes<sup>21,22</sup>. Therefore, in these tissues, *hth* can function without interference from the Hox genes.

These results suggest that *exd* and *hth* generate antennal identities. Because a nuclear-localized form of Exd was unable to trigger

**Figure 3** *Antp* represses *hth* expression and results in cytoplasmic Exd. **a–d**, An antennal disc with *Antp*-expressing cells. The disc was stained for Hth (magenta, **a**, **c**), Exd (green, **a**, **b**) and *Antp* (red, **a**, **d**); overlap between Hth and Exd appears white (**a**). A large clone of *Antp*<sup>+</sup> cells showing no Hth and cytoplasmic Exd is indicated by the white line. In proximal cells *Antp* expression does not lead to repression of *hth* (arrow). **e**, **f**, An *Antp*<sup>-</sup> clone in a leg disc stained for Exd (green). The clone was identified by the absence of CD2 staining (red) and is indicated by the white outline; the wild-type twin clone can be seen as the intense red area. In **f** only the Exd stain is shown. Although most of the *Antp*<sup>-</sup> cells have nuclear Exd (arrowhead), some cells have cytoplasmic Exd (arrow). **g**, **h**, A wild-type leg (**g**) and a leg with an *Antp*<sup>-</sup> clone, identified by the absence of  $\gamma$ <sup>+</sup> (**h**). Some *Antp*<sup>-</sup> cells develop into a third antennal segment (arrowhead), while others develop into normal leg structures, as indicated by the bracted  $\gamma$ <sup>-</sup> bristle (arrow). This heterogeneity can be explained if *Antp* protein persists long enough after the generation of the *Antp*<sup>-</sup> cell to stably repress *hth* in some of its descendants.



**Figure 4** *hth/Meis1* directs antennal development. **a**, The anal primordium of a genital disc expressing Myc-Meis1 in the *Dll* domain stained for Exd (green) and Myc-Meis1 (magenta); colocalization appears white. Both proteins are nuclear in the *Dll* domain. **b**, **c**, Wild-type anal plates (**b**) and a pair of anal plates transformed into antennae by expressing Myc-Meis1 in the *Dll* domain (**c**). The transformation includes a third antennal segment (arrow) and arista (arrowhead). Similar transformations result from ectopic Hth expression (not shown). **d**, Two truncated T2 legs (arrows) resulting from green fluorescent protein (GFP)-Hth expression in the *Dll* domain. **e**, **f**, A T3 leg disc (L, top) and a lateral view of a haltere disc (H, bottom) in which GFP-Hth was expressed in the *Dll* domain. The discs were stained for GFP-Hth (blue) and *Antp* (red); colocalization appears white; in **f** only the *Antp* stain is shown. Although GFP-Hth maintains high levels of *Antp* expression in the leg disc (arrowhead), it does not induce *Antp* expression *de novo* in the haltere disc (arrow).



antennal development, we suggest that *hth* contributes to antennal development in more ways than by localizing Exd to the nucleus. Because Hth contains a highly conserved homeodomain<sup>10</sup>, it, like Exd, is likely to be a DNA-binding transcription factor. However, although we favour the idea that *hth* and *exd* both contribute to antennal development, we cannot distinguish unambiguously between the roles of these two genes.

In leg discs, *Antp* prevents antennal development by repressing *hth*, returning the pathway to the ground state, which is leg. Although *Antp* plays no part in generating this ground state, it could contribute to normal leg development by modifying the ground state. We further suggest that repression of *hth* by *Antp* is linked to the growth of the leg disc, as is the formation of the proximodistal axis<sup>23</sup>. This is consistent with a previous study<sup>24</sup> showing that early pulses of *Antp* expression during antennal development tend to transform more distal structures, whereas later pulses tend to transform more proximal structures. Finally, our results suggest that *hth/exd* act positively to maintain *Antp* expression in leg discs. A previous study has shown *exd* to be a positive regulator of *Ubx* expression in imaginal discs<sup>25</sup>. In the case described here, this genetic circuitry appears to result in a feedback loop in which *Antp* expression is gradually lost during leg development because it represses its own transcriptional activator, *hth*.

Our results support the view<sup>1,2,26</sup> that the ancestral insect possessed leg-like appendages in head and thoracic segments. These primitive legs could have developed without *Antp*, *hth* or *exd*. Later, *hth* might have promoted the modification of this primitive limb into antenna. *Antp*, by repressing *hth*, returns the pathway to the ground state, allowing leg development to proceed in the thorax.

**Note added in proof:** Pai *et al.* also report an antenna to leg transformation in *hth*<sup>-</sup> clones<sup>31</sup>. □

## Methods

All alleles are from the Bloomington Stock Center or have been described<sup>10</sup>. *Antp*<sup>-</sup> or *hth*<sup>-</sup> clones were generated 24–48 h and *Antp*<sup>+</sup> flip-out clones 48–72 h after egg laying. For mitotic and flip-out clones, absence of the CD2 marker was used to identify mutant cells. A constitutively nuclear form of Exd was generated by fusing Exd to a nuclear form of  $\beta$ -galactosidase (*NLS-lacZ-exd*) (M. Abu-Shaar and R.S.M., in preparation). *NLS- $\beta$ -gal-Exd* is nuclear in the absence of *hth* and, when expressed in the *Dll* domain, it generates leg truncations that are similar to those produced by wild-type Exd<sup>20</sup>. *hth*<sup>-</sup> clones: *y HS-FLP; FRT82 hth<sup>Cl</sup>/FRT82 CD2 y<sup>+</sup>; Dll-lacZ* refers to *Dll*<sup>01092</sup>; dominant *Antp* allele: *Antp<sup>73b</sup>/TM3; Antp*<sup>-</sup> clones: *y HS-FLP; FRT82 Antp<sup>Ns+RC3</sup>/FRT82 CD2 y<sup>+</sup>; Antp<sup>+</sup> flip-out clones<sup>17,27</sup>; y HS-FLP; actin P > CD2 > Gal4; UAS-Antp; Dll-Gal4* expression<sup>19</sup>; *Dll-Gal4; UAS-Myc-Meis*, *UAS-GFP-Hth* or *UAS-NLS-lacZ-exd*. The anal plate to antenna transformation was most penetrant when *Dll-Gal4; UAS-Myc-Meis-8* larvae were grown at 17°C for four days and then shifted to 25°C. Antibodies used were: rabbit anti-Exd<sup>10</sup>; mouse anti-*Antp* (mAb 8C11)<sup>15</sup>; mouse anti-CD2 (Serotec); mouse anti-*Dll*<sup>28</sup> anti- $\beta$ -galactosidase (Promega); mouse anti-Myc (Oncogene Research). GFP was visualized directly and a chicken anti-Hth antibody was generated (Cocalico) against a histidine-tagged Hth fusion protein (residues 44 (an *SphI* site) to 232 (a *SacI* site)) expressed from pQE (Qiagen).

Received 17 December 1997; accepted 24 February 1998.

- Lewis, E. B. A gene complex controlling segmentation in *Drosophila*. *Nature* **276**, 565–570 (1978).
- Akam, M. E., Dawson, I. & Tear, G. Homeotic genes and the control of segment diversity. *Development* (suppl.) **123**–134 (1988).
- Gehring, W. Bildung eines vollständigen Mittelbeiner mit Sternopleura in der Antennenuregion bei der Mutante *Nasobemia* (Ns) von *Drosophila melanogaster*. *Arch. Klaus-Stift. Vererb.Forsch.* **41**, 44–54 (1966).
- Struhl, G. A homeotic mutation transforming leg to antenna in *Drosophila*. *Nature* **292**, 635–638 (1981).
- Struhl, G. Genes controlling segmental specification in the *Drosophila* thorax. *Proc. Natl Acad. Sci. USA* **79**, 7380–7384 (1982).
- Abbot, M. K. & Kaufman, T. C. The relationship between the functional complexity and the molecular organization of the *Antennapedia* locus of *Drosophila melanogaster*. *Genetics* **114**, 919–942 (1986).
- Schnewly, S., Klemenz, R. & Gehring, W. J. Redesigning the body plan of *Drosophila* by ectopic expression of the homeotic gene *Antennapedia*. *Nature* **325**, 816–818 (1987).
- Gonzalez-Crespo, S. & Morata, G. Control of *Drosophila* adult pattern by *extradenticle*. *Development* **121**, 2117–2125 (1995).

- Rauskolb, C., Smith, K., Peifer, M. & Wieschaus, E. *extradenticle* determines segmental identities throughout development. *Development* **121**, 3663–3671 (1995).
- Rieckhof, G., Casares, F., Ryoo, H. D., Abu-Shaar, M. & Mann, R. S. Nuclear translocation of *Extradenticle* requires *homothorax*, which encodes an *Extradenticle*-related homeodomain protein. *Cell* **91**, 171–183 (1997).
- Kaufman, T. C., Lewis, R. & Wakimoto, B. T. Cytogenetic analysis of chromosome 3 in *Drosophila melanogaster*: the homeotic gene complex in polytene chromosome interval 84A-B. *Genetics* **94**, 115–133 (1980).
- Schnewly, S., Kuroiwa, A. & Gehring, W. J. Molecular analysis of the dominant homeotic *Antennapedia* phenotype. *EMBO J.* **6**, 201–206 (1987).
- Carroll, S. B., Laymon, R. A., McCutcheon, M. A., Riley, P. D. & Scott, M. P. The localization and regulation of *Antennapedia* protein expression in *Drosophila* embryos. *Cell* **47**, 113–122 (1986).
- Wirz, J., Fessler, L. I. & Gehring, W. J. Localization of the *Antennapedia* protein in *Drosophila* embryos and imaginal discs. *EMBO J.* **5**, 3327–3334 (1986).
- Condie, J. M., Mustard, J. A. & Brower, D. Generation of anti-*Antennapedia* monoclonal antibodies and *Antennapedia* protein expression in imaginal discs. *Drosoph. Inf. Serv.* **70**, 52–54 (1991).
- Rauskolb, C., Peifer, M. & Wieschaus, E. *extradenticle*, a regulator of homeotic gene activity, is a homolog of the homeobox-containing human proto-oncogene *pbx1*. *Cell* **74**, 1–20 (1993).
- Struhl, G. & Basler, K. Organizing activity of wingless protein in *Drosophila*. *Cell* **72**, 527–540 (1993).
- Brand, A. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118**, 401–415 (1993).
- Calleja, M., Moreno, E., Pelaz, S. & Morata, G. Visualization of gene expression in living adult *Drosophila*. *Science* **274**, 252–255 (1996).
- Gonzalez-Crespo, S. & Morata, G. Genetic evidence for the subdivision of the arthropod limb into coxopodite and telopodite. *Development* **122**, 3921–3928 (1996).
- Freeland, D. E. & Kuhn, D. T. Expression patterns of developmental genes reveal segment and parasegment organization of *D. melanogaster* genital discs. *Mech. Dev.* **56**, 61–72 (1996).
- Casares, F., Sánchez, L., Guerrero, I. & Sánchez-Herrero, E. The genital disc of *Drosophila melanogaster*. I. Segmental and compartmental organization. *Dev. Genes Evol.* **207**, 216–228 (1997).
- Lecuit, T. & Cohen, S. M. Proximal–distal axis formation in the *Drosophila* leg. *Nature* **388**, 139–145 (1997).
- Gibson, G. & Gehring, W. J. Head and thoracic transformation caused by ectopic expression of *Antennapedia* during *Drosophila* development. *Development* **102**, 657–675 (1988).
- Azpiazu, N. & Morata, G. Functional and regulatory interactions between *Hox* and *extradenticle* genes. *Genes Dev.* **12**, 261–273 (1998).
- Shubin, N., Tabin, C. & Carroll, S. Fossils, genes and the evolution of animal limbs. *Nature* **388**, 639–648 (1997).
- Pignoni, F. & Zipursky, S. L. Induction of *Drosophila* eye development by decapentaplegic. *Development* **124**, 271–278 (1997).
- Panganiban, G., Sebring, A., Nagy, L. & Carroll, S. The development of crustacean limbs and the evolution of arthropods. *Science* **270**, 1363–1366 (1995).
- Cohen, S. M. Specification of limb development in the *Drosophila* embryo by positional cues from segmentation genes. *Nature* **343**, 173–177 (1990).
- Steiner, E. Establishment of compartments in the developing leg imaginal discs of *Drosophila melanogaster*. *Wilhelm Roux Arch. Dev. Biol.* **180**, 9–30 (1976).
- Pai, C.-Y. *et al.* The homothorax homeoprotein activates the nuclear localization of another homeoprotein, *Extradenticle*, and suppresses eye development in *Drosophila*. *Genes Dev.* **12**, 435–446 (1998).

**Acknowledgements.** We thank G. Struhl for fly stocks (including the FRT *Antp*<sup>Ns+RC3</sup> chromosome) and for generating a new set of *Antp*<sup>M</sup> clones; H. D. Ryoo for preparing the Hth protein for making the anti-Hth antibody; M. Abu-Shaar for generating the NLS–Exd fusion; G. Rieckhof for first characterizing *hth*<sup>-</sup> clones; G. Panganiban and D. Brower for antibodies; and G. Morata, E. Sánchez-Herrero, G. Struhl and members of our laboratory for discussions and comments on the manuscript. This work was supported by grants from the HFSP and NIH to R.S.M. E.C. was supported by an EMBO fellowship and R.S.M. is a scholar of the Leukemia Society of America.

Correspondence and requests for materials should be addressed to R.S.M. (e-mail: rsm10@columbia.edu).

## Forced degradation of Fas inhibits apoptosis in adenovirus-infected cells

Ann E. Tollefson\*, Terry W. Hermiston\*†, Drew L. Lichtenstein\*, Clarence F. Collier\*, Ralph A. Tripp‡, Todor Dimitrov\*, Karoly Toth\*, Christopher E. Wells\*, Peter C. Doherty‡ & William S. M. Wold\*

\* Department of Molecular Microbiology and Immunology, St Louis University School of Medicine, 1402 South Grand Boulevard, St Louis, Missouri 63104-1004, USA

‡ Department of Immunology, St Jude Children's Research Hospital, 332 North Lauderdale, Memphis, Tennessee 38105-2794, USA

DNA viruses have evolved elaborate mechanisms to overcome host antiviral defences. In adenovirus-infected cells, programmed cell death (apoptosis) induced by the cytokine tumour necrosis factor (TNF) is inhibited by several adenovirus-encoded proteins<sup>1–3</sup>. Occupation of the cell-surface receptor Fas, a member of the TNF-receptor superfamily that is expressed on most cell types,

† Present address: Onyx Pharmaceuticals, 3031 Research Drive, Richmond, California 94806, USA.

triggers apoptosis of that cell<sup>4-6</sup>. Here we show that the adenovirus RID (for receptor internalization and degradation) protein complex, which is an inhibitor of TNF-induced apoptosis<sup>2</sup>, mediates internalization of cell-surface Fas and its destruction inside lysosomes within the cell. Fas has not previously been shown to be internalized and then degraded. RID also mediates internalization of the receptor for epidermal growth factor<sup>7,8</sup>, but it does not affect the transferrin receptor or class I antigens of the major histocompatibility complex. Removal of Fas from the surface of adenovirus-infected cells expressing RID may allow infected cells to resist Fas-mediated cell death and thus promote their survival.

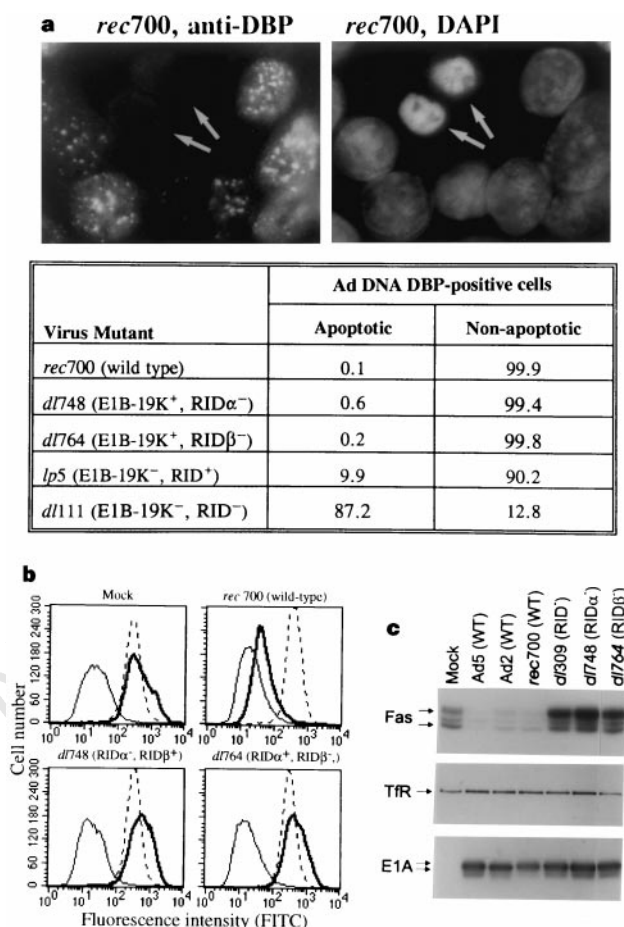
RID is composed of two adenovirus E3-region-encoded polypeptides with relative molecular masses of 10.4K and 14.5K (refs 9, 10), named RID- $\alpha$  and RID- $\beta$  respectively. The adenovirus E1B-19K protein also inhibits TNF-induced apoptosis<sup>3</sup>. To determine whether these proteins inhibit Fas-mediated apoptosis, cells were infected with wild-type adenovirus or with adenovirus mutants that were unable to express the RID- $\alpha$ , RID- $\beta$ , or E1B-19K proteins. Figure 1a shows a list of the mutants and the proteins

they lack. Cells were treated with a monoclonal antibody that triggers Fas-induced apoptosis, and then were fixed and stained with 4,6-diamidino-2-phenylindole (DAPI) to visualize nuclear DNA and with antibody to the adenovirus-encoded DNA-binding protein (DBP) to identify infected cells. Most cells infected with wild-type adenovirus (named *rec700*) (such cells are indicated by the speckled staining for DBP in the nuclei) had non-apoptotic nuclei (Fig. 1a). Two uninfected cells had apoptotic nuclei (arrows). Only ~1% of cells infected with wild-type adenovirus or with mutants that lacked RID- $\alpha$  (*dl748* virus) or RID- $\beta$  (*dl764* virus) were apoptotic (Fig. 1a). About 10% of the cells infected with a mutant lacking E1B-19K (*lp5* virus) were apoptotic. However, when a mutant lacking RID- $\alpha$ , RID- $\beta$ , and E1B-19K (*dl111* virus) was used, ~90% of the cells were apoptotic. Thus adenovirus expresses two proteins that independently inhibit Fas-induced apoptosis, RID (that is, RID- $\alpha$  and RID- $\beta$ ) and E1B-19K. E1B-19K also inhibits Fas-induced apoptosis in E1B-19K-transfected cells<sup>11</sup>.

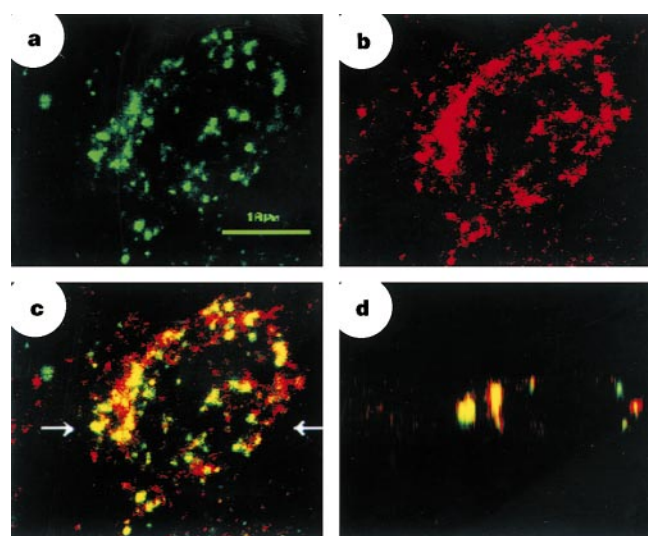
The E1B-19K protein is a homologue of Bcl-2<sup>12,13</sup>, an inhibitor of Fas-mediated apoptosis, which may explain why E1B-19K inhibits apoptosis. RID is a membrane-located cell-surface protein<sup>9,10,14,15</sup> that stimulates internalization and degradation of the epidermal growth factor (EGF) receptor<sup>7,8</sup>. RID could exert a similar effect on Fas and thereby inhibit apoptosis. Indeed, most Fas protein was cleared from cells infected with wild-type adenovirus (Fig. 1b). The transferrin receptor was not affected, indicating that receptor clearance is not a general feature in adenovirus-infected cells. Fas was not cleared from cells infected with mutants lacking RID- $\alpha$  or RID- $\beta$ . Thus, RID is necessary to clear Fas from the cell surface. Similar results were obtained in other cells (A549 cells; results not shown).

Many receptors are internalized into endosomes and degraded in lysosomes. Fas was degraded in adenovirus-infected cells if both RID- $\alpha$  and RID- $\beta$  were present (Fig. 1c). The transferrin receptor was not degraded. The reduction in Fas protein levels was not due to lack of Fas messenger RNA (results not shown). All infections were equivalent as indicated by expression of the adenovirus E1A proteins (Fig 1c).

Immunofluorescence study of wild-type adenovirus-infected cells showed that Fas was present in numerous vesicles and not on the cell surface (results not shown). When mutant viruses lacking RID- $\alpha$  or RID- $\beta$  were used, Fas was found at the plasma membrane. To determine whether some of the vesicles were lysosomes, wild-type



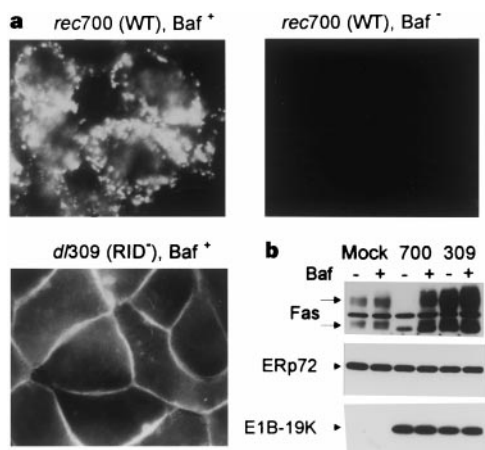
**Figure 1** RID is required for inhibition of apoptosis induced by Fas agonist, internalization of cell-surface Fas, and degradation of Fas in MCF7-Fas cells infected with adenovirus mutants. **a**, Cells were infected with *rec700*, treated with an agonist monoclonal antibody against Fas, then double-stained for the adenovirus-encoded DBP and for DNA. Infected cells with speckled staining of DBP (left) had non-apoptotic DAPI-stained nuclei (right). Arrows indicate uninfected cells with apoptotic nuclei. The table shows the percentage of apoptotic nuclei in mutant-infected cells. Ad, adenovirus. **b**, Cells were infected with viruses, incubated with antibodies against Fas (bold plot) or the transferrin receptor (dashed plot) or with control IgG (thin plot) then analysed by flow cytometry. Mock, mock-infected cells. **c**, Proteins from infected cells were analysed by immunoblotting for Fas, transferrin receptor (TfR) or adenovirus E1A protein. WT, wild-type.



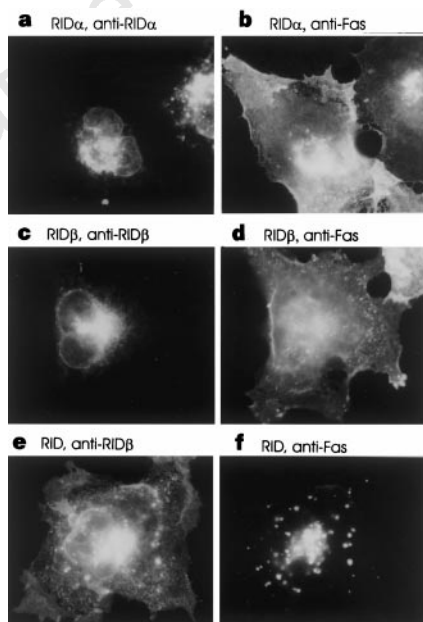
**Figure 2** Fas colocalizes with LAMP-1 in lysosomes in *rec700*-infected cells. **a**, Cells labelled with rabbit anti-Fas antibody and FITC. **b**, Cells labelled with mouse anti-LAMP1 antibody and RITC. **c**, Combined image of **a**, **b**. **d**, Perpendicular view of image in **c** (arrows), 1  $\mu$ m thick. Yellow indicates colocalization. Bar, 10  $\mu$ m.

adenovirus-infected cells were double-stained for Fas and lysosome-associated membrane protein (LAMP-1, a lysosomal protein<sup>16</sup>) and examined by confocal microscopy (Fig. 2). Green, red, and yellow vesicles contain Fas (Fig. 2a), LAMP-1 (Fig. 2b), or both Fas and LAMP-1 (Fig. 2c, d), respectively. The many yellow vesicles seen shows that Fas colocalizes with LAMP-1 in lysosomes. The Fas-containing green vesicles may be endosomes. Similar results were obtained when using another lysosomal protein, CD63 (results not shown).

Bafilomycin A1 (Baf) specifically inhibits the vacuolar-type H<sup>+</sup>-ATPase, preventing the acidification of vesicles and the trafficking of receptors from late endosomes to lysosomes<sup>17,18</sup>. When cells infected with wild-type adenovirus were treated with Baf, Fas was cleared from the cell surface but it accumulated in vesicles rather than being degraded as in untreated cells (Fig. 3a). Baf did not reduce cell-surface Fas in cells infected with a mutant lacking RID (*d/309* virus). When examined by immunoblotting, Baf could be seen to block



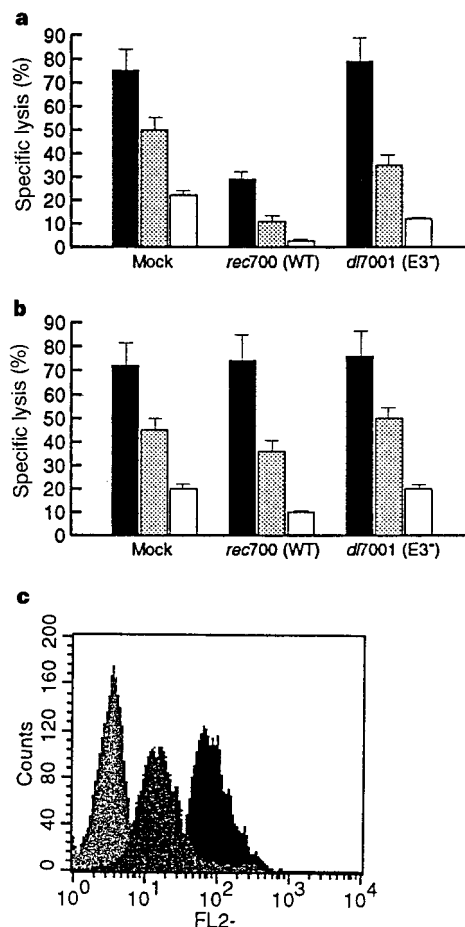
**Figure 3** Bafilomycin A1 (Baf) inhibits RID-mediated degradation of Fas. **a**, Immunofluorescence labelling of Fas in *rec700*-infected cells treated (top left) or not treated (top right) with Baf, or *d/309* (RID<sup>-</sup>)-infected cells treated with Baf (bottom left). **b**, Immunoblot of Fas, ERp72, or adenovirus protein E1B-19K in mock-, *rec700*- or *d/309*-infected cells. WT, wild-type.



**Figure 4** The presence of RID is sufficient to internalize cell-surface Fas into vesicles in COS cells transiently cotransfected with RID-α, RID-β, and Fas. Cells were transfected with RID-α plus Fas (**a**, **b**), RID-β plus Fas (**c**, **d**), or RID-α, RID-β, and Fas (**e**, **f**), and then double-stained for RID-α (**a**), RID-β (**c**, **e**), and Fas (**b**, **d**, **f**). **a** and **b**, **c** and **d**, and **e** and **f** are the same fields.

the degradation of Fas in cells infected with wild-type adenovirus (Fig. 3b). Baf did not affect the abundance of Fas in mock-infected cells or in cells infected with the mutant virus lacking RID. Fas levels were actually increased when this mutant was used; this observation is not understood. Neither virus infection nor Baf treatment affected the abundance of ERp72, a cellular protein that is localized in the endoplasmic reticulum<sup>19</sup>. The infections were equivalent as indicated by the E1B-19K levels (Fig. 3b).

These results indicate that RID is needed for the internalization of Fas into lysosomes (presumably via endosomes), the degradation of Fas in lysosomes, and the inhibition of Fas-induced apoptosis. To determine whether the presence of RID is sufficient to induce the internalization of Fas, COS cells were transiently cotransfected with combinations of Fas, RID-α, and RID-β, and then were double-stained for RID-α and Fas, or for RID-β and Fas. With RID-α plus Fas, or RID-β plus Fas, Fas was localized on the cell surface (Fig. 4b, d). However, in cells triple-transfected with RID-α, RID-β, and Fas, Fas was found in vesicles rather than on the cell surface (Fig. 4f). When both RID-α and RID-β were present, RID-β staining was localized to the endoplasmic reticulum and plasma membrane (Fig. 4e), probable sites of RID action<sup>14,15</sup>. Inhibition of Fas-induced apoptosis was studied in MCF7-Fas cells that were transiently cotransfected with plasmid vectors expressing RID-α or RID-β, treated with the agonist monoclonal antibody to Fas, then fixed in acetone/DAPI and stained for RID-β. When RID-β-expressing



**Figure 5** RID is needed to inhibit killing of adenovirus-infected P815 cells by CTLs from perforin<sup>-/-</sup> mice. Effector-lymphocyte:target ratios were 60:1 (black bars), 20:1 (stippled bars), or 6:1 (open bars). **a**, Results obtained with perforin<sup>-/-</sup> lymphocytes. **b**, Results obtained with perforin<sup>+/+</sup> lymphocytes. **c**, Flow cytometry of P815 cells infected with *rec700* (wild-type; middle plot) or *d/7000* (RID<sup>-</sup>; right plot). The left plot shows the IgG control. WT, wild-type. FL2-, fluorescence intensity.



vectors alone were used, 80% of RID- $\beta$ -positive cells had apoptotic nuclei. In contrast, when vectors expressing both RID- $\alpha$  and RID- $\beta$  were used, only 26% of RID- $\beta$ -positive cells had apoptotic nuclei. With the vector alone, 62% of the total cells were apoptotic. Thus, RID (that is, RID- $\alpha$  plus RID- $\beta$ ), but not RID- $\beta$  or RID- $\alpha$  alone, is sufficient for internalization of Fas and inhibition of Fas-induced apoptosis.

Cytotoxic T lymphocytes (CTLs) carrying the antigen CD8 induce apoptosis through two main pathways, one using perforin/granzymes and the other using Fas<sup>20</sup>. To examine CTL-mediated killing through Fas, a short-term CD3-dependent redirected cell assay<sup>21</sup> was performed, using CTLs from perforin<sup>-/-</sup> mice<sup>22</sup> and adenovirus-infected P815 cells. The perforin<sup>-/-</sup> CTLs lysed mock-infected cells efficiently (Fig. 5a). Lysis was inhibited by wild-type adenovirus but not by a mutant lacking all E3 genes. CTLs from perforin<sup>+/+</sup> mice killed mock-infected cells and cells infected with wild-type or E3 mutant adenovirus with similar high efficiency (Fig. 5b). Cell-surface Fas levels were decreased on cells infected with wild-type adenovirus but not on cells infected with an E3 mutant (Fig. 5c). Thus, an E3 protein, presumably RID, inhibits CTL-mediated killing through the Fas pathway by reducing Fas levels on the cell surface.

RID may prolong acute adenovirus infections by clearing Fas from the cell surface and reducing killing by leukocytes such as CTLs that express Fas ligands. RID could also promote persistent adenovirus infections in leukocytes. After virus infections, Fas ligand expressed on activated T cells induces apoptosis of leukocytes, thereby allowing the leukocyte population to return to steady-state levels<sup>4,23-25</sup>. Leukocytes infected with adenovirus might not be eliminated because RID has degraded Fas.

**Note added in proof:** Results analogous to those in Fig. 1a, b have recently been reported (Shisler, J., Yang, C., Walter, B., Ware, C. F., & Gooding, L. R. The adenovirus E3-10.4K/14.5K complex mediates loss of cell surface Fas (CD95) and resistance to Fas-induced apoptosis. *J. Virol.* **71**, 8299-8306, (1997)). □

## Methods

**Cells and viruses.** All cells used for adenovirus infections were MCF7-Fas cells, the MCF7 human cell line that is stably transfected with Fas<sup>26</sup>, except in the experiments used to obtain the results in Fig. 5, where mouse P815 cells were used. Adenovirus 5/2rec700 (refs 9, 10), adenovirus 5 and adenovirus 2 are wild-type viruses. *dl748* and *dl764* virus mutants lack RID- $\alpha$  and RID- $\beta$ , respectively<sup>9,10</sup>; *lp5* lacks E1B-19K but has intact E3 (ref. 27); *dl309* lacks RID- $\alpha$  and RID- $\beta$ ; and *dl111* lacks E1B-19K, RID- $\alpha$ , and RID- $\beta$ <sup>28</sup>. *dl7001* lacks all E3 proteins and *dl7000* lacks all E3 proteins except E3-14.7K (ref. 29).

**Apoptosis assay.** Cells were infected with 250 plaque-forming units (PFU) per cell of virus, except that 10 PFU of *lp5* and *dl111* were used. At 21 h after infection, cells were treated for 22 h with CH-11, an agonist monoclonal antibody to Fas (200 ng ml<sup>-1</sup>; Panvera), plus cycloheximide (25  $\mu$ g ml<sup>-1</sup>). Cells were fixed in methanol containing DAPI and stained for DBP using a rabbit antiserum and goat anti-rabbit IgG/fluorescein isothiocyanate (FITC). The percentage of apoptotic nuclei was calculated (at least 1,000 DBP-positive cells were examined per sample). Apoptotic nuclei were shrunken and irregular, had condensed DNA, and often fluoresced very brightly above the plane of focus for non-apoptotic nuclei.

For transfections, the RID- $\alpha$  and RID- $\beta$  genes were cloned into the pMT2 plasmid vector to generate pMT2-RID- $\alpha$  and pMT2-RID- $\beta$ . pcDNA3-Fas expresses Fas<sup>26</sup>. MCF7-Fas cells were cotransfected with pMT2-RID- $\beta$  plus pMT2, or with pMT2-RID- $\beta$  plus pMT2-RID- $\alpha$  (2.5  $\mu$ g for each plasmid). After 38 h, cells were treated for 9 h with the CH-11 monoclonal antibody to Fas (500 ng ml<sup>-1</sup>) plus cycloheximide, fixed in methanol/DAPI and stained for RID- $\beta$ <sup>15</sup>; 100-200 RID- $\beta$ -positive cells per sample were scored for apoptotic nuclei.

**Immunofluorescence.** Virus-infected cells were fixed using 3.7% paraformaldehyde followed by methanol/DAPI. Cells were double-stained with a rabbit anti-Fas antibody (Santa Cruz Biotechnology) and the BB6 mouse anti-human-LAMP-1 monoclonal antibody<sup>16</sup>, followed by goat anti-rabbit IgG-

FITC and goat anti-mouse IgG-RITC (Cappel ICN). Cells were examined using a Zeiss LSM 410 scanning laser confocal microscope with LSM 410 software. C057 cells were transfected with 1  $\mu$ g of each plasmid DNA. After 30 h, cells were fixed in methanol/DAPI and double-stained for Fas (using the ZB4 monoclonal antibody; Panvera) and for RID- $\alpha$  and RID- $\beta$ <sup>15</sup>.

**Flow cytometry and immunoblotting.** Cells were infected with adenovirus mutants. At 27 h after infection, cells were suspended in FACS buffer containing mouse IgG $\gamma$  (5  $\mu$ g ml<sup>-1</sup>), anti-human-Fas UB2 IgG monoclonal antibody (10  $\mu$ g ml<sup>-1</sup>; Panvera), or antibody against human transferrin receptor (2.5  $\mu$ g ml<sup>-1</sup>; Boehringer/Mannheim), incubated with goat anti-mouse FITC-conjugated antibody (20  $\mu$ g ml<sup>-1</sup>), and analysed on a FACScaliber flow cytometer using Cell Quest software. For immunoblots, rabbit antisera to Fas (Santa Cruz), ERp72 (ref. 19), or E1B-19K were used, as were anti-transferrin-receptor monoclonal antibody OKT9 (ATCC) or anti-E1A monoclonal antibody M73, followed by the appropriate peroxidase-conjugated secondary antibody. Proteins were detected with ECL reagents. In Fig. 5, the clone Jo2 anti-mouse monoclonal antibody was used.

**Bafilomycin treatment.** Infected cells were treated with Baf (0.1  $\mu$ M) for 12 h beginning at 13 h after infection, then immunostained for Fas. For immunoblot analysis, cells were treated with Baf at 6 h after infection and processed 18 h later.

**CTL cytotoxicity assay.** Splenocytes were isolated from perforin<sup>-/-</sup> mice<sup>22</sup> and matched wild-type C57BL/6J H-2<sup>b</sup> mice primed with influenza A virus<sup>30</sup>. CTLs were generated by incubation with irradiated influenza-virus-infected splenocytes, then with the 145-2C11 anti-CD3 $\epsilon$  monoclonal antibody for 30 min. P815 cells were infected with 1,000 PFU per cell of adenovirus mutant, labelled with Na<sub>2</sub><sup>51</sup>CrO<sub>4</sub>, and incubated with the activated anti-CD3 $\epsilon$ -treated CTLs. Cell lysis was determined 6 h later by release of <sup>51</sup>Cr.

Received 22 September 1997; accepted 15 January 1998.

- Gooding, L. R., Elmore, L. W., Tollefson, A. E., Brady, H. A. & Wold, W. S. M. A 14,700 MW protein from the E3 region of adenovirus inhibits cytolysis by tumor necrosis factor. *Cell* **53**, 341-346 (1988).
- Gooding, L. R. *et al.* The 10,400- and 14,500-dalton proteins encoded by region E3 of adenovirus function together to protect many but not all mouse cell lines against lysis by tumor necrosis factor. *J. Virol.* **65**, 4114-4123 (1991).
- Gooding, L. R. *et al.* The E1B 19,000-molecular-weight protein of group C adenoviruses prevents tumor necrosis factor cytolysis of human cells but not of mouse cells. *J. Virol.* **65**, 3083-3094 (1991).
- Nagata, S. Apoptosis by death factor. *Cell* **88**, 355-365 (1997).
- Itoh, N. *et al.* The polypeptide encoded by the cDNA for human cell surface antigen Fas can mediate apoptosis. *Cell* **66**, 233-243 (1991).
- Oehm, A. *et al.* Purification and molecular cloning of the APO-1 cell surface antigen, a member of the tumor necrosis factor/nerve growth factor receptor superfamily. *J. Biol. Chem.* **267**, 10709-10715 (1992).
- Carlin, C. R., Tollefson, A. E., Brady, H. A., Hoffman, B. L. & Wold, W. S. M. Epidermal growth factor receptor is down-regulated by a 10,400 MW protein encoded by the E3 region of adenovirus. *Cell* **57**, 135-144 (1989).
- Tollefson, A. E., Stewart, A. R., Yei, S. P., Saha, S. K. & Wold, W. S. M. The 10,400- and 14,500-dalton proteins encoded by region E3 of adenovirus form a complex and function together to down-regulate the epidermal growth factor receptor. *J. Virol.* **65**, 3095-3105 (1991).
- Tollefson, A. E., Krajcsi, P., Yei, S. P., Carlin, C. R. & Wold, W. S. M. A 10,400-molecular-weight membrane protein is coded by region E3 of adenovirus. *J. Virol.* **64**, 794-801 (1990).
- Tollefson, A. E., Krajcsi, P., Pursley, M. H., Gooding, L. R. & Wold, W. S. M. A 14,500 MW protein is coded by region E3 of group C human adenoviruses. *Virology* **175**, 19-29 (1990).
- Hashimoto, S., Ishii, A. & Yonehara, S. The E1B oncogene of adenovirus confers cellular resistance to cytotoxicity of tumor necrosis factor and monoclonal anti-Fas antibody. *Int. Immunol.* **3**, 343-351 (1991).
- Boyd, J. M. *et al.* Adenovirus E1B 19 kDa and Bcl-2 proteins interact with a common set of cellular proteins. *Cell* **79**, 341-351 (1994).
- White, E. Life, death, and the pursuit of apoptosis. *Genes Dev.* **10**, 1-15 (1996).
- Hoffman, P., Yaffe, M. B., Hoffman, B. L., Yei, S., Wold, W. S. M. & Carlin, C. Characterization of the adenovirus E3 protein that down-regulates the epidermal growth factor receptor. Evidence for intermolecular disulfide bonding and plasma membrane localization. *J. Biol. Chem.* **267**, 13480-13487 (1992).
- Stewart, A. R., Tollefson, A. E., Krajcsi, P., Yei, S. P. & Wold, W. S. M. The adenovirus E3 10.4K and 14.5K proteins, which function to prevent cytolysis by tumor necrosis factor and to down-regulate the epidermal growth factor receptor, are localized in the plasma membrane. *J. Virol.* **69**, 172-181 (1995).
- Carlsson, S. R., Roth, J., Piller, F. & Fukuda, M. Isolation and characterization of human lysosomal membrane glycoproteins, h-lamp-1 and h-lamp-2. Major sialoglycoproteins carrying poly(lactosaminoglycan). *J. Biol. Chem.* **15**, 18911-18919 (1988).
- Yoshimori, T., Yamamoto, A., Moriyama, Y., Futai, M. & Tashiro, Y. Bafilomycin A1, a specific inhibitor of vacuolar-type H<sup>+</sup>-ATPase, inhibits acidification and protein degradation in lysosomes of cultured cells. *J. Biol. Chem.* **266**, 17707-17712 (1991).
- van Weert, A. W. M., Dunn, K. W., Geuze, H. J., Maxfield, F. R. & Stoorvogel, W. Transport from late endosomes to lysosomes, but not sorting of integral membrane proteins in endosomes, depends on the vacuolar proton pump. *J. Cell. Biol.* **130**, 821-834 (1995).
- Mazzarella, R. A., Srinivasan, M., Haugejorden, S. & Green, M. ERp72, an abundant, luminal ER protein, contains three copies of the active site sequences of protein disulfide isomerase. *J. Biol. Chem.* **265**, 1094-1101 (1990).
- Clark, W. R. *et al.* Molecular pathways of CTL-mediated cytotoxicity. *Immunol. Rev.* **146**, 33-44 (1995).
- Azuma, M., Cayabyab, M., Buck, D., Phillips, J. H. & Lanier, L. L. CD28 interaction with B7

- costimulates primary allogeneic proliferative responses and cytotoxicity mediated by small, resting T lymphocytes. *J. Exp. Med.* 175, 353–360 (1992).
22. Kagi, D. *et al.* Fas and perforin pathways as major mechanisms of T cell-mediated cytotoxicity. *Science* 265, 528–530 (1994).
  23. Brunner, T. *et al.* Cell-autonomous Fas (CD95)/Fas-ligand interaction mediates activation-induced apoptosis in T-cell hybridomas. *Nature* 373, 441–444 (1995).
  24. Dhein, J., Walczak, H., Baumler, C., Debatin, K.-M. & Krammer, P. H. Autocrine T-cell suicide mediated by APO-1/(Fas/CD95). *Nature* 373, 438–441 (1995).
  25. Ju, S.-T. *et al.* Fas(CD95)/Fas.L interactions required for programmed cell death after T-cell activation. *Nature* 373, 444–448 (1995).
  26. Jäättelä, M., Benedict, M., Tewari, M., Shayman, J. A. & Dixit, V. M. Bcl-x and Bcl-2 inhibit TNF and Fas-induced apoptosis and activation of phospholipase A<sub>2</sub> in breast carcinoma cells. *Oncogene* 10, 2297–2305 (1995).
  27. Subramanian, T., Kuppuswamy, M., Mak, S. & Chinnadurai, G. Adenovirus *cyt* locus, which controls cell transformation and tumorigenicity, is an allele of *lp* locus, which codes for a 19-kilodalton tumor antigen. *J. Virol.* 52, 336–343 (1984).
  28. Babiss, L. E., Fisher, P. B. & Ginsberg, H. S. Effect on transformation of mutations in the early region 1b-encoded 21- and 55-kilodalton proteins of adenovirus 5. *J. Virol.* 52, 389–395 (1984).
  29. Ranheim, T. S. *et al.* Characterization of mutants within the gene for the adenovirus E3 14.7-kilodalton protein which prevents cytolysis by tumor necrosis factor. *J. Virol.* 67, 2159–2167 (1993).
  30. Tripp, R. A., Hou, S., McMickle, A., Houston, J. & Doherty, P. C. Recruitment and proliferation of CD8<sup>+</sup> T cells in respiratory virus infections. *J. Immunol.* 154, 6013–6021 (1995).

**Acknowledgements.** We thank J. Freeman for help with the confocal microscopy, V. Dixit for the MCF7 and MCF7-Fas cell lines and for pcDNA-3-Fas, M. Green for the pMT2 vector, the ERp72 antiserum, and COS7 cells, M. Green for the DBP antiserum, M. Fukuda for LAMP1 antiserum, E. Harlow for the M73 mAb, C. Gugliemio for technical assistance, J. Mikes for preparation of the manuscript and figures, and C. Pollack for photography.

Correspondence and requests for materials should be addressed to W.S.M.W. (e-mail: woldws@slu.edu).

## Integrin binding and mechanical tension induce movement of mRNA and ribosomes to focal adhesions

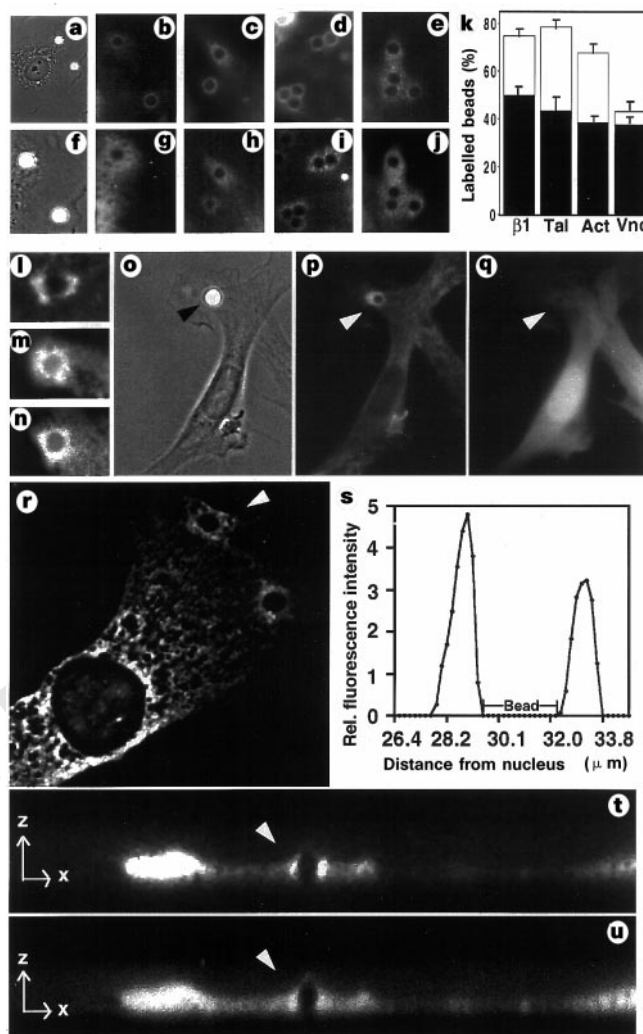
Marina E. Chicurel<sup>†</sup>, Robert H. Singer<sup>†</sup>, Christian J. Meyer<sup>\*</sup> & Donald E. Ingber<sup>\*</sup>

<sup>\*</sup> Departments of Surgery and Pathology, Children's Hospital and Harvard Medical School, 300 Longwood Avenue, Boston, Massachusetts 02115, USA

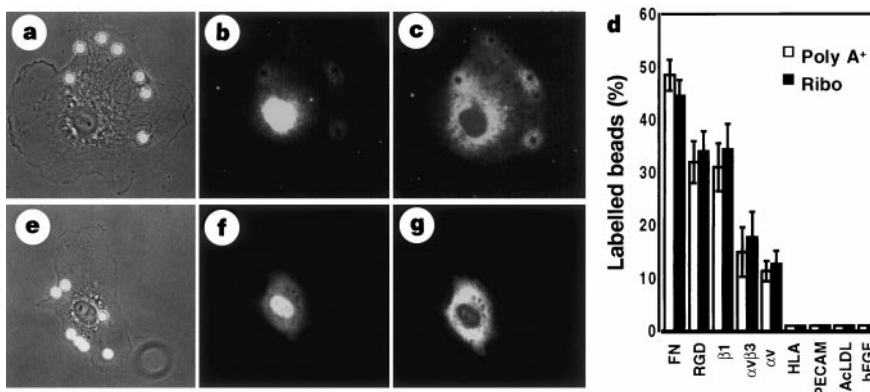
<sup>†</sup> Department of Anatomy and Structural Biology, Albert Einstein College of Medicine of Yeshiva University, Bronx, New York 10461, USA

The extracellular matrix (ECM) activates signalling pathways that control cell behaviour by binding to cell-surface integrin receptors and inducing the formation of focal adhesion complexes (FACs)<sup>1,2</sup>. In addition to clustered integrins, FACs contain proteins that mechanically couple the integrins to the cytoskeleton<sup>3</sup> and to immobilized signal-transducing molecules<sup>1,2</sup>. Cell adhesion to the ECM also induces a rapid increase in the translation of pre-existing messenger RNAs<sup>4,5</sup>. Gene expression can be controlled locally by targeting mRNAs to specialized cytoskeletal domains<sup>6</sup>. Here we investigate whether cell binding to the ECM promotes formation of a cytoskeletal microcompartment specialized for translational control at the site of integrin binding. High-resolution *in situ* hybridization revealed that mRNA and ribosomes rapidly and specifically localized to FACs that form when cells bind to ECM-coated microbeads. Relocation of these protein synthesis components to the FAC depended on the ability of integrins to mechanically couple the ECM to the contractile cytoskeleton and on associated tension-moulding of the actin lattice. Our results suggest a new type of gene regulation by integrins and by mechanical stress which may involve translation of mRNAs into proteins near the sites of signal reception.

FACs containing  $\beta 1$  integrin, talin, actin and vinculin could be visualized along the bead–membrane interface 20 min after human umbilical vein endothelial cells had bound to microbeads (4.5  $\mu\text{m}$ ) coated with fibronectin (FN) (Fig. 1a–f)<sup>2,7</sup>. High-resolution *in situ* hybridization using oligonucleotide probes for poly(A)<sup>+</sup> and ribosomal RNAs revealed that both mRNA and ribosomes were simultaneously recruited to the region surrounding the FAC (Fig. 1g–j). Whereas most FAC-associated cytoskeleton (CSK) proteins were



**Figure 1** Recruitment of mRNA and ribosomes to FACs 20 min after cell binding to ECM-coated microbeads. Low- (a) and high- (f) magnification phase-contrast micrographs of a cell 20 min after binding to FN-coated beads. Fluorescence views showing immunostaining for  $\beta 1$  integrin (b) and *in situ* hybridization for ribosomal RNA (g) in the same cell as shown in f; bead diameter, 4.5  $\mu\text{m}$ . Immunofluorescence staining for other FAC proteins, including talin (c), actin (d), and vinculin (e), in similarly treated cells. Corresponding *in situ* hybridization results obtained in the same cells using probes for poly(A)<sup>+</sup> (h, i) or ribosomal RNA (j) are shown below. k, Graph showing the percentage of beads that exhibited positive staining for each of the FAC proteins (white bars) and the subset of these beads that also stained positive for either poly(A)<sup>+</sup> RNA or ribosomes (black bars).  $\beta 1$ ,  $\beta 1$  integrin; Tal, talin; Act, actin; Vnc, vinculin. Error bars indicate s.e.m. Similar recruitment of mRNA was observed in human microvascular lung endothelial cells (l), bovine capillary endothelial cells (m), and NIH3T3 fibroblasts (n) bound to RGD-beads, as detected using an FITC-labelled oligo(dT) probe. Phase contrast (o) and corresponding fluorescence (p, q) views of a 3T3 fibroblast expressing high levels of soluble cytoplasmic enhanced green fluorescent protein through transfection. Binding to an RGD-coated bead (arrowheads) specifically induced recruitment of ribosomes (p) but not the green fluorescent protein (q). r, Digital imaging optical section in the x, y (horizontal) plane showing fluorescence staining for ribosomes in an endothelial cell bound to two FN-coated microbeads. s, Graph showing the relative fluorescence intensity in a single optical section due to ribosomal staining in the region of the cytoplasm surrounding one of the FN-coated beads shown in r (arrowhead). The average relative fluorescence intensity in a nearby cytoplasmic region at an equal distance from the nucleus in the same cell, but without an attached bead, was  $\sim 0.57 \pm 0.67$  in this study. Confocal micrographs in the x, z (vertical) plane showing fluorescence staining for poly(A)<sup>+</sup> RNA (t) and ribosomes (u) in a spread endothelial cell bound to a FN-coated microbead. Arrowheads indicate increased densities of poly(A)<sup>+</sup> RNA and ribosomes at the site of bead binding.

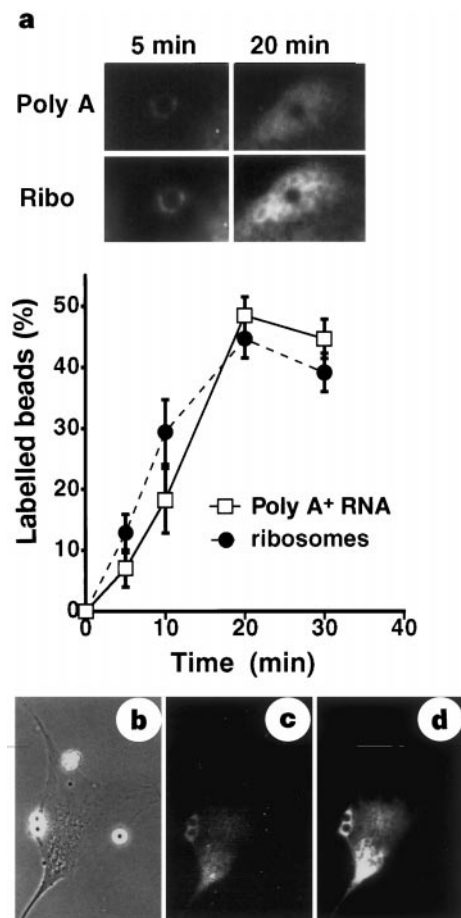


**Figure 2** Integrin-dependent relocation of poly(A)<sup>+</sup> RNA and ribosomes to the FAC at 20 min after bead binding. Phase contrast (**a**, **e**) and corresponding fluorescence (**b**, **c**, **f**, **g**) images of cells bound to beads coated with antibodies against integrin  $\beta 1$  (**a**–**c**) or FGF protein (**e**–**g**) and stained with an FITC-labelled oligo(dT) probe (**b**, **f**) or a Cy3-labelled ribosomal probe (**c**, **g**). Note that the FGF-coated beads that are incapable of inducing FAC formation show no bead-associated staining. **d**, Quantification of the percentage of beads that stained positively for poly(A)<sup>+</sup> RNA and ribosomes within cells bound to beads coated with FN (FN); RGD peptide (RGD); antibodies against  $\beta 1$  integrin ( $\beta 1$ ),  $\alpha \nu \beta 3$  integrin ( $\alpha \nu \beta 3$ ),  $\alpha \nu$  integrin ( $\alpha \nu$ ), HLA histocompatibility antigen (HLA), or PECAM; or with the metabolic receptor ligands, acetylated low-density lipoprotein (AcLDL) or FGF.

restricted to the bead–membrane interface, mRNA and ribosomes appeared to form larger amorphous haloes that encompassed the FAC and extended farther into the cytoplasm ( $\sim 2.5 \mu\text{m}$ ) at this time. Vinculin gave a similar diffuse staining pattern (Fig. 1e, j). Quantification of staining at this time (20 min after bead binding) revealed that only a subset of beads recruited FAC proteins<sup>2,7</sup> and that only a subset of these FACs stained positively for mRNA or ribosomes (Fig. 1k). However, beads labelled for mRNA almost invariably contained FAC proteins. Similar results were obtained using beads coated with RGD peptide and with other cell types (Fig. 1l–n). In contrast, specific staining was not observed when fixed cells were pretreated with RNase before hybridization or when a ‘sense’ oligo(dA) probe was used instead of oligo(dT). Furthermore, soluble cytoplasmic components, such as transfected green fluorescent protein (GFP), did not appear to accumulate around bead-induced FACs (Fig. 1o–q). Digital imaging microscopy in the *x*, *y* plane (Fig. 1r, s) and confocal microscopy in the *x*, *z* plane (Fig. 1t, u) confirmed that the local densities of ribosomes and poly(A)<sup>+</sup> RNA were indeed specifically increased in the region of the bead-associated FAC.

Different integrin subtypes differed in their ability to promote poly(A)<sup>+</sup> RNA and ribosome redistribution. Beads coated with anti-integrin  $\beta 1$  antibodies (Fig. 2a–c) induced more mRNA and ribosome relocation than did beads coated with either anti- $\alpha \nu \beta 3$  or anti- $\alpha \nu$  antibodies (Fig. 2d). In contrast, beads coated with ligands for other (non-integrin) transmembrane receptors, including specific antibodies (against HLA antigen, or PECAM) or metabolic receptor ligands, such as basic fibroblast growth factor (FGF) or acetylated low-density lipoprotein, failed to induce any visible recruitment (Fig. 2d–g). Thus, the observed relocation of mRNA and ribosomes in response to ECM binding appeared to be specifically mediated by integrin binding and associated FAC formation.

In migrating fibroblasts,  $\beta$ -actin mRNA localizes to the CSK at the leading edge within a few minutes of growth factor stimulation<sup>8</sup>. Similarly, staining for poly(A)<sup>+</sup> RNA and ribosomes could be detected delineating the bead–membrane interface within the FAC 5 min after bead binding; the staining pattern only later became more diffuse and extended into the surrounding cytoplasm (Fig. 3a). The percentage of positively stained beads also progressively increased over time, saturating at  $\sim 20$  min after bead binding (Fig. 3a). These rapid kinetics suggested that transcription and nuclear transport might not be required for RNA localization to FACs. To test this, we enucleated cells to produce cytoplasts and then allowed them to bind to FN-coated beads (Fig. 3b–d). The percentage of beads that exhibited recruitment of poly(A)<sup>+</sup> RNA and ribosomes was nearly identical in nucleated cytoplasts and



**Figure 3** Kinetics of mRNA and ribosome recruitment to the FAC. **a**, Top, fluorescence staining for poly(A)<sup>+</sup> RNA (Poly A) and ribosomes (Ribo) around a cell-surface-bound FN-coated bead at 5 and 20 min after binding. Graph shows time course of recruitment of poly(A)<sup>+</sup> RNA and ribosomes to the FAC. Phase-contrast (**b**) and corresponding fluorescence views (**c**, **d**) of an enucleated cytoplast bound to FN-coated beads that stained positively for poly(A)<sup>+</sup> RNA (**c**) and ribosomes (**d**).

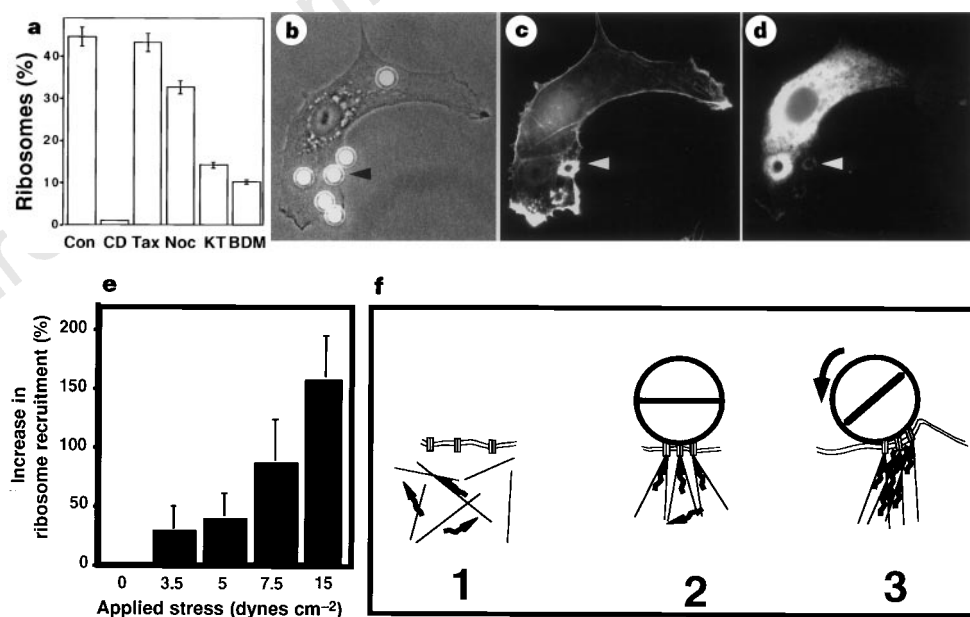


control cells ( $34.25 \pm 4.97$  compared with  $36.63 \pm 1.92$ , respectively). Thus, neither *de novo* synthesis of RNA nor vectorial transport of RNA from the nucleus to the cytoplasm was required for the repositioning of poly(A)<sup>+</sup> RNA and ribosomes to the FAC.

Compartmentalization of specific mRNAs observed in many cell types, including oocytes, neurons, muscle cells and fibroblasts, appears to be mediated by the targeting of existing cytoplasmic mRNAs to specific microdomains of the insoluble CSK<sup>6</sup>. In fact, disruption of the actin lattice using cytochalasin D blocked recruitment of poly(A)<sup>+</sup> RNA and ribosomes to the FAC in our cells (Fig. 4a). In contrast, treatment of cells with microtubule-modifying drugs, such as nocodazole and taxol, had only minimal effects. Poly(A)<sup>+</sup> RNA could therefore be passively recruited to the FAC owing to its association with actin filaments<sup>9</sup>, which accumulate at the bead surface as the cell extends ruffles around the bead in response to integrin binding (Fig. 4b, c). We consider this unlikely, however, because beads that stained positively for actin were negative or very dimly stained for poly(A)<sup>+</sup> RNA and ribosomes and vice versa (Fig. 4b–d).

Another possibility is that this microcompartmentalization is controlled mechanically through tension-dependent restructuring of the CSK. For example, certain mRNAs preferentially associate with vertices in the actin lattice and not with the intervening struts<sup>10</sup>. These fine microarchitectural features can be generated through changes in isometric tension resulting from cell binding to an ECM substrate that can resist cell-tractional forces<sup>11</sup>. If tension-moulding of the CSK is required for the observed redistribution of the protein synthesis machinery, then changing the balance of forces transmitted across the cell surface, for example by altering mechanical coupling between integrins and the CSK, should modify this recruitment response. We found that there was a direct correlation

between the ability of transmembrane receptors to cause mRNA and ribosomes to relocate to the FAC and their ability to resist mechanical distortion, as previously demonstrated using magnetic twisting cytometry<sup>12</sup>. Beads that formed the strongest CSK attachments (coated with RGD peptide or anti-β1 integrin antibodies) were consistently more effective at inducing recruitment than beads coated with receptor ligands which gave intermediate (anti-αvβ3 and anti-αv antibodies) or low (AcLDL, anti-HLA) levels of mechanical coupling to the CSK (Fig. 2). Furthermore, treatment of cells with drugs such as KT5926 and butanedione-2-monoxime (BDM) that inhibit actomyosin-based tension generation without disrupting CSK filaments<sup>13,14</sup> also decreased recruitment of poly(A)<sup>+</sup> RNA and ribosomes to the FAC (Fig. 4a). Finally, we could directly demonstrate that tension *per se* is a trigger for these repositioning events by applying mechanical stresses to cell surface integrin receptors using controlled magnetic twisting forces and cell magnetometry<sup>3</sup>. Application of mechanical stresses ranging from 0 to 15 dynes cm<sup>-2</sup> to cell-surface-bound magnetic beads coated with RGD peptide resulted in a stress-dependent increase in the percentage of beads that stained for ribosomes after 5 min stress application (Fig. 4e). Myosin V is required for the transport of *ASH1* mRNA in budding yeast<sup>15</sup> and localization of β-actin mRNA to the leading edge appears to involve the small GTP-binding protein Rho (V. Latham and R.H.S., unpublished results), which promotes FAC formation by increasing myosin light-chain phosphorylation and increasing CSK tension<sup>16</sup>. These results indicate that cell adhesion to the ECM and integrin binding may create a specialized microcompartment containing mRNA and ribosomes near the FAC through mechanical restructuring of the surrounding actin lattice and the corresponding formation of geometries that promote mRNA and ribosome binding (Fig. 4f).



**Figure 4** The role of the CSK and mechanical tension in the recruitment process. **a**, Percentage of beads that recruited ribosomes in control untreated cells (Con) or in cells treated for 30 min with 1  $\mu\text{g ml}^{-1}$  cytochalasin D (CD), 1  $\mu\text{M}$  Taxol (Tax), or 10  $\mu\text{g ml}^{-1}$  nocodazole (Noc) or for 1 h with 15  $\mu\text{M}$  KT5926 (KT) or 20 mM BDM. Phase contrast (**b**) and corresponding fluorescence views (**c**, **d**) of the same cell bound to FN-coated beads stained for F-actin using FITC-phalloidin (**c**) or ribosomes (**d**). Arrowhead indicates a bead that is brightly stained for F-actin but only dimly stained for ribosomes, whereas the adjacent bead to the left is only weakly stained with phalloidin but brightly stained for ribosomes. **e**, Mechanical-stress-dependent recruitment of ribosomes to the FAC. Rotational shear stresses ranging from 0 to 15 dynes cm<sup>-2</sup> were applied to cell-surface-bound magnetic beads coated with RGD peptide using controlled magnetic twisting forces<sup>3</sup>. Error bars indicate s.e.m. **f**, Model of tension-dependent recruitment of mRNA and

ribosomes to the FAC. Integrin receptors (rectangles) on the cell membrane (1) are induced to cluster and form a FAC linked to the actin CSK (black lines) in response to binding to a FN-coated bead (large circle) (2). The ability of the bound bead to resist CSK tension results in a low level of CSK reorganization that exposes sites that promote binding of ribosomes and mRNA (arrows) previously present within the cytoplasm (2). Increasing the level of mechanical distortion within the actin lattice by magnetically twisting the bound beads results in progressive tension-moulding of the lattice, creation of additional binding sites, and thus enhanced recruitment of mRNA and ribosomes to the site of integrin binding (3). This restructuring and microcompartmentalization starts at the bead-membrane interface and then progressively extends outwards into the cytoplasm, as in Fig. 3.

Our results suggest that rapid post-transcriptional changes in gene expression may be mediated by repositioning of translational components to sites of signal reception and that mechanical tension may guide these movements. FACs contain several signal-transducing molecules, including proteins that can regulate translation (the  $\text{Na}^+/\text{H}^+$  antiporter<sup>1,17,18</sup>, protein kinase C<sup>19–21</sup>, PI-3 kinase<sup>1,2,22</sup> and the FGF receptor<sup>1,23</sup>, for example). Thus, redistribution of mRNA and ribosomes to the FAC may explain the rapid increase in protein synthesis that is observed well before changes in transcription are detectable in response to substrate adhesion<sup>4</sup>. Other FAC components, such as calreticulin<sup>24</sup> and proteins containing LIM domains (such as zyxin, cysteine-rich protein-1 and paxillin)<sup>25</sup>, can potentially bind nucleic acids and may thus also play a role in the observed response to stress and integrin binding. Formation of this localized complex could also mediate the post-transcriptional effects of mechanical stresses on gene expression. For example, in cardiocytes, an acute rise in ventricular wall stress results in a rapid increase in protein synthesis and phosphorylation of the translation initiation factor eIF-4E, effects that can be prevented by treatment with BDM<sup>26</sup>. These observations suggest that translation of gene products at the site of ECM signal reception may constitute a novel aspect of integrin signalling and serve an intermediate role between rapid local post-translational changes (such as tyrosine phosphorylation, local CSK rearrangements) and slower transcriptional alterations in the nucleus. Formation of this microcompartment for post-transcriptional control of gene expression also may serve to integrate signals that regulate cellular protein synthesis with those that are elicited by integrins, growth-factor receptors and mechanical stresses within the same FAC<sup>1–3</sup>. □

## Methods

**Cell culture.** Endothelial cells from human umbilical vein<sup>27</sup>, bovine adrenal cortex<sup>17</sup>, or human lung microvasculature (Clonetics) and NIH3T3 cells were cultured on FN-coated glass coverslips in serum-containing medium for 14–16 h before bead addition. NIH3T3 cells were transfected with a plasmid encoding an enhanced green fluorescent protein (Clontech) using lipofectamine (GIBCO-BRL). Cytoplasts were generated<sup>28</sup> and incubated with beads 3 h after cytochalasin B washout. Polystyrene beads (4.5  $\mu\text{m}$ ; Polysciences) were coated with FN (Cappel), AcLDL (Biomedical Technologies), FGF (Takeda Chemical), RGD-containing peptide (Peptide 2000; Telios) or rabbit anti-mouse IgG (Cappel) as described<sup>29</sup>. Beads coated with anti-IgG were incubated with purified antibodies (25  $\mu\text{g ml}^{-1}$ ) to HLA (clone W 6-32, ATCC), PECAM (clone P2B1, Chemicon), or integrin types  $\beta 1$  (clone B-D15, Biosource),  $\alpha \nu \beta 3$  (clone LM609, Chemicon), or  $\alpha \nu$  (clone P3G8, Chemicon). Beads were added to cells (30 beads per cell) in serum-free medium with 1% BSA, washed free of unattached beads, and fixed in paraformaldehyde. Uncoated beads, beads coated only with anti-mouse IgG antibodies, or beads coated with antibodies directed against intracellular proteins did not bind to cell surfaces.

**In situ hybridization.** *In situ* hybridization was performed using probes directly labelled with either FITC or Cy3 fluorochromes as described<sup>30</sup>. An oligo(dT) 43-mer was used to assess poly(A)<sup>+</sup> RNA distribution; an oligo(dA) 43-mer was used as a negative control. Ribosomes were detected using a mixture of 4 oligonucleotide probes complementary to the human 18S ribosomal RNA (all 5' to 3': ACGGTATCTGATCGTCTTCGAACCTCCGACT; ACCCGCACTTACTGGAATTCCTCGTTCAT; TTATTCGTCATCTCGGGTGGCTGAACGCCACT; GCCTCACTAACCATCCAATCGGTAGTAGC). An anti-ribosomal antibody produced similar results. Hybridizations were carried out at 37 °C for 2 h in 10% (oligo(dT) probes) or 30% formamide (ribosomal probes). The highest stringency washes were performed at 37 °C in 1 $\times$  SSC, 10% formamide (for oligo(dT) probes) or 30% formamide (for ribosomal probes). Quantitative data are expressed as the pooled average of more than 230 beads obtained from at least two independent experiments.

**Immunofluorescence.** Immunofluorescence microscopy was performed as described<sup>7</sup> using antibodies at the following dilutions: anti- $\beta 1$  integrin (clone B-D15, Biosource) 1:40, anti-talin (Chemicon) 1:25, anti-vinculin (clone hVIN-1, Sigma) 1:400, and anti-ribosomal P antigen (Immunovision) 1:50, FITC-labelled sheep anti-mouse IgG (Amersham) 1:50, Cy3-labelled goat anti-

mouse IgG (Amersham) 1:1,000, and biotin-labelled goat anti-human IgG (Vector) 1:200. FITC- or Texas red-labelled avidin D (Vector) was used at a dilution of 1:250 and FITC-labelled phalloidin (Sigma) at 1:100. A Leica TCS4D confocal microscope was used to obtain optical sections in the vertical ( $x, z$ ) plane. Digital (deconvolution) imaging microscopy was used to obtain horizontal ( $x, y$ ) optical sections and to quantify relative fluorescence intensities<sup>9</sup>.

**Magnetic twisting cytometry.** Mechanical stresses were applied directly to cell-surface integrin receptors by magnetically twisting surface-bound ferromagnetic beads (4.5  $\mu\text{m}$ ; SpheroTech) that were coated with RGD peptide<sup>3</sup>. A 1,000 gauss magnetic field was first applied for 10  $\mu\text{s}$  to align the moments of the beads 5 min after binding to the cell surface and then 0–30 gauss magnetic fields were applied in an orthogonal direction for 5 more minutes to exert a controlled shear stress (torque) to the beads bound to cell-surface receptors.

Received 24 September 1997; accepted 3 February 1998.

- Plopper, G., McNamee, H. P., Dike, L. E., Bojanowski, K. & Ingber, D. Convergence of integrin and growth factor receptor signaling pathways within the focal adhesion complex. *Mol. Biol. Cell.* **6**, 1349–1365 (1995).
- Miyamoto, S. *et al.* Integrin function: molecular hierarchies of CSK and signaling molecules. *J. Cell Biol.* **131**, 791–805 (1995).
- Wang, N. J., Butler, P. & Ingber, D. E. Mechanotransduction across the cell surface and through the CSK. *Science* **260**, 1124–1127 (1993).
- Benecke, B.-J., Ben-Ze'ev, A., Benecke, B.-J. & Penman, S. The control of mRNA production, translation and turnover in suspended and reattached anchorage-dependent fibroblasts. *Cell* **14**, 931–939 (1978).
- Farmer, S. R., Ben-Ze'ev, A. & Penman, S. Altered translatability of messenger RNA from suspended anchorage-dependent fibroblasts: reversal upon cell attachment to a surface. *Cell* **15**, 627–637 (1978).
- Bassell, G. & Singer, R. H. mRNA and cytoskeletal filaments. *Curr. Opin. Cell Biol.* **9**, 109–115 (1997).
- Plopper, G. & Ingber, D. E. Rapid induction and isolation of focal adhesion complexes. *Biochem. Biophys. Res. Commun.* **193**, 571–578 (1993).
- Latham, V., Kislaukas, E., Singer, R. & Ross, A.  $\beta$ -Actin mRNA localization is regulated by signal transduction mechanisms. *J. Cell Biol.* **126**, 1211–1219 (1994).
- Taneja, K. L., Lifshitz, L. M., Fay, F. S. & Singer, R. H. Poly(A) RNA codistribution with microfilaments: Evaluation by *in situ* hybridization and quantitative digital imaging microscopy. *J. Cell Biol.* **119**, 1245–1260 (1992).
- Bassell, G. J., Powers, C. M., Taneja, K. L. & Singer, R. H. Single mRNAs visualized by ultrastructural *in situ* hybridization are principally localized at actin filament intersections in fibroblasts. *J. Cell Biol.* **126**, 863–876 (1994).
- Ingber, D. E. Cellular tensegrity: defining new rules of biological design that govern the cytoskeleton. *J. Cell Sci.* **104**, 613–627 (1993).
- Wang, N. & Ingber, D. E. Probing transmembrane mechanical coupling and cytomechanics using magnetic twisting cytometry. *Biochem. Cell Biol.* **73**, 327–335 (1995).
- Cramer, L. P. & Mitchison, T. J. Myosin is involved in postmitotic cell spreading. *J. Cell Biol.* **131**, 179–189 (1995).
- Nakanishi, S., Yamada, K., Iwahashi, K., Kuroda, K. & Kase, H. KT5926, a potent and selective inhibitor of myosin light chain kinase. *Mol. Pharmacol.* **37**, 482–488 (1990).
- Long, R. M. *et al.* Mating type switching in yeast controlled by asymmetric localization of ASH1 mRNA. *Science* **277**, 383–387 (1997).
- Chrzanowska-Wodnicka, M. & Burridge, K. Rho-stimulated contractility drives the formation of stress fibers and focal adhesions. *J. Cell Biol.* **133**, 1403–1415 (1996).
- Ingber, D. E. *et al.* Control of intracellular pH and growth by fibronectin in capillary endothelial cells. *J. Cell Biol.* **110**, 1803–1811 (1990).
- Edmonds, B. T., Murray, J. & Condeelis, J. pH regulation of the F-actin binding properties of Dictyostelium elongation factor 1a. *J. Biol. Chem.* **270**, 15222–15230 (1995).
- Vuori, K. & Ruoslahti, E. Activation of protein kinase C precedes  $\alpha 5 \beta 1$  integrin-mediated cell spreading on FN. *J. Biol. Chem.* **268**, 21459–21462 (1993).
- Morley, S. J. Signal transduction mechanisms in the regulation of protein synthesis. *Mol. Biol. Rep.* **19**, 221–231 (1994).
- Whalen, S. G. *et al.* Phosphorylation of eIF-4E on serine 209 by protein kinase C is inhibited by the translational repressors, 4E-binding proteins. *J. Biol. Chem.* **271**, 11831–11837 (1996).
- Von Manteuffel, S., Gingras, A. C., Ming, X. F., Sonenberg, N. & Thomas, G. 4E-BP1 phosphorylation is mediated by the FRAP-p70s6k pathway and is independent of mitogen-activated protein kinase. *Proc. Natl Acad. Sci. USA* **93**, 4076–4080 (1996).
- Sonenberg, N. Regulation of translation and cell growth by eIF-4E. *Biochimie* **76**, 839–846 (1994).
- Singh, N. K., Attreya, C. D. & Nakhasi, H. L. Identification of calreticulin as a rubella virus RNA binding protein. *Proc. Natl Acad. Sci. USA* **91**, 12770–12774 (1994).
- Schmeichel, K. L. & Beckerle, M. C. Molecular dissection of a LIM domain. *Mol. Biol. Cell.* **8**, 219–230 (1997).
- Wada, H., Ivester, C. T., Carabello, B. A., Cooper, G. & McDermott, P. J. Translational initiation factor eIF-4E: A link between cardiac load and protein synthesis. *J. Biol. Chem.* **271**, 8359–8364 (1996).
- Kourembanas, S., Hannan, R. L. & Faller, D. V. Oxygen tension regulates the expression of the platelet-derived growth factor-B chain in human endothelial cells. *J. Clin. Invest.* **86**, 670–674 (1990).
- Lucas, J. J., Szekeley, E. & Kates, J. R. The regeneration and division of mouse L-cell karyoplasts. *Cell* **7**, 115–122 (1976).
- Schwartz, M. A., Lechene, C. & Ingber, D. E. Insoluble FN activates the  $\text{Na}^+/\text{H}^+$  antiporter by clustering and immobilizing integrin  $\alpha 5 \beta 1$ , independent of cell shape. *Proc. Natl Acad. Sci. USA* **88**, 7849–7853 (1991).
- Zhang, G., Taneja, K. L., Singer, R. H. & Green, M. R. Localization of pre-mRNA splicing in mammalian nuclei. *Nature* **372**, 809–812 (1994).

**Acknowledgements.** We thank K. Taneja, S. Shenoy, N. Wang, J. Fredberg, S. Kourembanas, G. Lee, G. Whitesides and the Biomedical Imaging Facility at the University of Massachusetts in Worcester for key reagents, equipment and technical expertise, and B. Peters and C. Chen for discussion. This work was supported by grants from the NIH and NASA and an American Cancer Society postdoctoral fellowship.

Correspondence and requests for materials should be addressed to D.E.I. (e-mail: ingber@al.tch.harvard.edu).

# Electron currents generated by the human phagocyte NADPH oxidase

Jacques Schrenzel<sup>\*†</sup>, Lena Serrander<sup>\*</sup>, Botond Bánfi<sup>\*‡</sup>, Oliver Nüße<sup>\*</sup>, Reyhaneh Fouyouzi<sup>\*</sup>, Daniel P. Lew<sup>\*</sup>, Nicolas Demaurex<sup>§</sup> & Karl-Heinz Krause<sup>\*</sup>

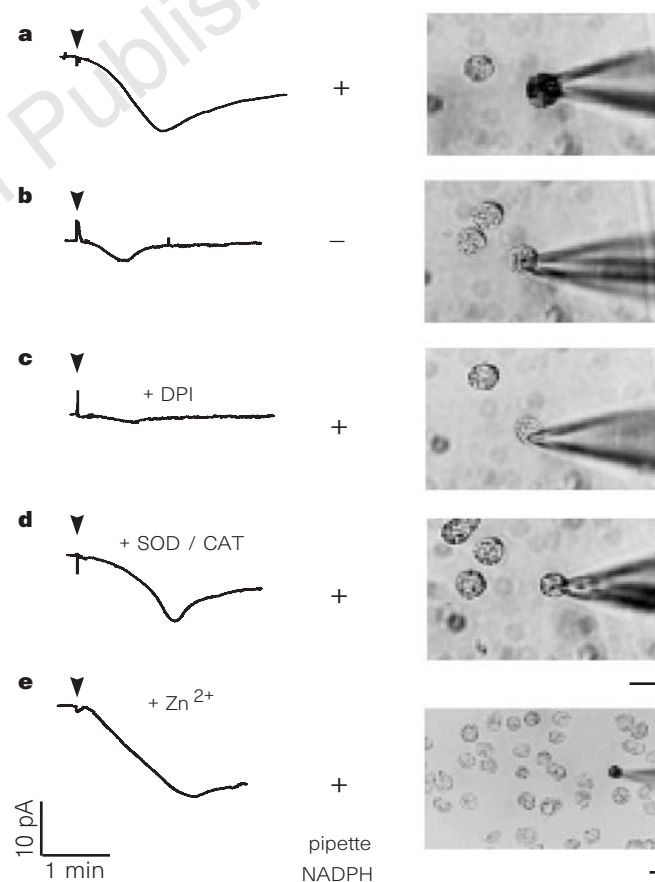
<sup>\*</sup> Division of Infectious Diseases, Department of Medicine, University Hospital, and <sup>§</sup> Department of Physiology, University Medical Center, 1211 Geneva 4, Switzerland

Electron transport across biological membranes is a well-known feature of bacteria, mitochondria and chloroplasts, where it provides motive forces for vectorial transport processes<sup>1</sup>. In contrast, electron transport is generally not found in the plasma membrane of eukaryotic cells, possibly because it would interfere with electric processes at the plasma membrane. An exception is provided by the phagocyte NADPH oxidase, which generates superoxide ( $O_2^{\cdot-}$ ) through electron transfer from cytosolic NADPH to extracellular oxygen<sup>2-5</sup>. The enzyme is essential for host defence, and patients with chronic granulomatous disease, who lack the functional enzyme, suffer from severe infections<sup>6,7</sup>. It has been suggested that electron transfer by the NADPH oxidase might be electrogenic<sup>8</sup>. Here we demonstrate, using the whole-cell patch-clamp technique, the generation of electron currents by the NADPH oxidase in human eosinophil granulocytes. The currents were absent in granulocytes of sufferers of chronic granulomatous disease and under conditions of low oxygen. Generation of electron currents across the plasma membrane of eukaryotic cells has not been observed previously and might be—independently of the generation of superoxide—a physiologically relevant function of the phagocyte NADPH oxidase.

On the basis of the amount of superoxide generated by granulocytes ( $\sim 10$  nmoles  $O_2^{\cdot-}$  per min per  $10^6$  cells<sup>9</sup>), the NADPH oxidase should transport  $\sim 10^8$  electrons per second per cell. If these electron fluxes occur through an electrogenic pathway, they would be expected to produce currents of up to 10–20 pA. Such currents might indeed play a role in the depolarization observed during granulocyte activation<sup>8,10-12</sup>. To identify putative currents mediated by NADPH oxidase, we analysed whole-cell currents in eosinophils under experimental conditions designed to minimize currents through known ion channels and transporters. In a first series of experiments, we used unbuffered, nominally  $Ca^{2+}$ -free solutions in the bath and pipette (that is, the free  $Ca^{2+}$  concentration was  $\sim 5$   $\mu$ M as determined by  $Ca^{2+}$ -sensitive electrodes) and included 8 mM NADPH in the pipette solution. Whole-cell patch-clamp recordings at a holding voltage of 0 mV showed the development of inward currents in 95% of eosinophils (Fig. 1a; peak current density:  $6.26 \pm 0.42$  pA/pF; cell capacitance:  $2.49 \pm 0.05$  pF;  $n = 64$ ). To determine whether the NADPH oxidase was activated under these conditions, we performed patch-clamp experiments with nitroblue tetrazolium (NBT) in the bath solution. NBT is reduced by reactive oxygen species to dark-blue, insoluble formazan–NBT (this represents positive NBT staining) and is widely used to detect activation of NADPH oxidase<sup>13</sup>. 94% of the patch-clamped cells, but none of the neighbouring unpatched cells, showed dark NBT precipitates (Fig. 1a). Thus, the inward currents were associated with the generation of oxygen radicals and their sign and amplitude was consistent with the predicted electron fluxes through the NADPH

oxidase. However, there could be a coincidental activation of NADPH-oxidase-unrelated currents, or the activation of currents through products of NADPH oxidase (reactive oxygen species).

Omission of NADPH from the pipette solution markedly decreased the size of the currents (peak current density:  $2.75 \pm 0.34$  pA/pF;  $n = 28$ ;  $P < 0.0001$ ) and there was no detectable NBT staining under these conditions (Fig. 1b). Cytosolic NADPH is important for the observed currents. The residual currents in the absence of pipette NADPH are probably due to endogenous NADPH, which is present in the cytosol directly after break-in but which is rapidly consumed and/or lost to the patch pipette during the course of the experiment. Pretreatment of cells with the NADPH oxidase inhibitor diphenyleneiodonium (DPI; ref. 14) strongly diminished the size of the inward currents (peak current density in DPI-pretreated cells:  $0.42 \pm 0.22$  pA/pF;  $n = 5$ ) and there was no NBT staining (Fig. 1c). In contrast, when superoxide dismutase and catalase were included in the bath solution, the inward currents were preserved (peak current density:  $5.11 \pm 0.76$  pA/pF;  $n = 10$ ), but no NBT staining was seen (Fig.



**Figure 1** Inward currents associated with activation of NADPH oxidase in human eosinophils. Whole-cell recordings from eosinophils<sup>15</sup> were made in the presence of 0.25 mM NBT to assay for activity of NADPH oxidase. After current recording (left panels), the cells were photographed to show the presence or absence of NBT staining (right panels). **a**, Pipette solution contained 8 mM NADPH; 15 out of 16 patched cells showed dark NBT precipitates. **b**, Pipette solution lacked NADPH; 0 out of 10 cells were NBT-positive. **c**, Cells were preincubated for 3 min with 10  $\mu$ M DPI, an inhibitor of NADPH oxidase; 0 out of 5 cells were NBT-positive. **d**, Bath solution contained superoxide dismutase (SOD, 50 units per ml) and catalase (CAT, 2,000 units per ml); 1 out of 10 cells were NBT-positive. **e**, Bath solution contained 10  $\mu$ M  $Zn^{2+}$  to inhibit  $H^+$  currents; 5 out of 5 cells were NBT-positive. Arrowheads mark the onset of the whole-cell recording (break-in); plus and minus signs indicate the presence or absence of NADPH (8 mM) in the pipette solution. Bars under right panels indicate 10  $\mu$ m.

<sup>†</sup> Present address: Division of Clinical Microbiology, Mayo Clinic, Rochester, 200 First Street SW, Minnesota 55905, USA.

<sup>‡</sup> Permanent address: Department of Physiology, Semmelweis Medical University, H-1444 Budapest 8, PO Box 259, Hungary.



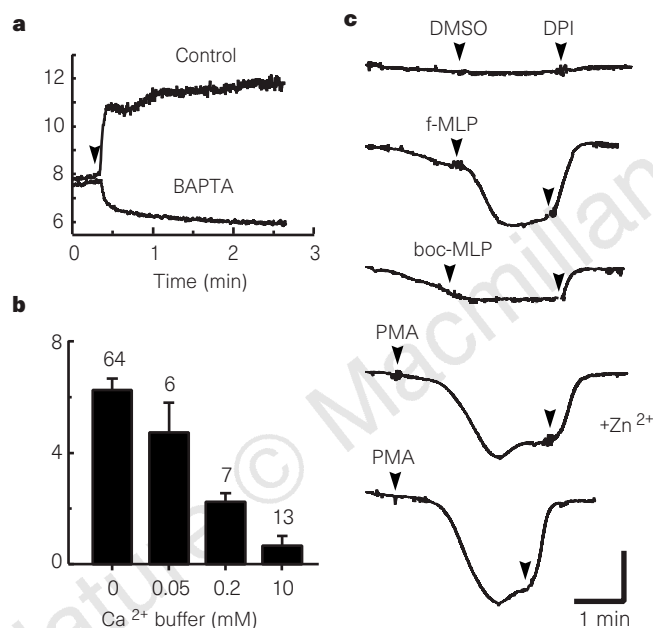
1d). Thus, superoxide dismutase and catalase removed the oxidase products without significantly affecting the currents.

To exclude components of  $H^+$  currents, we used  $10 \mu M$   $Zn^{2+}$  (a concentration that efficiently blocks eosinophil  $H^+$  currents<sup>15</sup>). As shown in Fig. 1e, current amplitude and kinetics were nearly identical to those seen in the absence of  $Zn^{2+}$  (peak current density:  $6.58 \pm 1.97$  pA/pF;  $n = 6$ ), excluding a relevant contribution of  $H^+$  currents to the whole-cell currents seen at 0 mV.

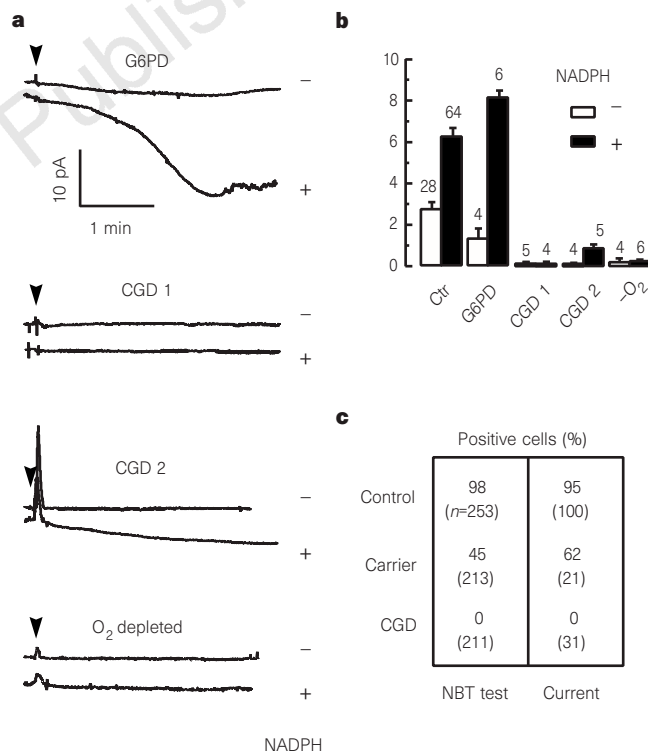
What led to activation of NADPH oxidase under our experimental conditions? As the experiments shown in Fig. 1 were performed using pipette solution without inclusion of  $Ca^{2+}$  buffers, the activation might be due to elevations of the cytosolic free  $[Ca^{2+}]_i$ . To investigate the possible role of  $Ca^{2+}$  in activation of NADPH oxidase in patch-clamped cells, we measured  $[Ca^{2+}]_i$  using the fluorescent dye indo-1. As shown in Fig. 2a, when using a nominally  $Ca^{2+}$ -free buffer,  $[Ca^{2+}]_i$  elevations occurred (the indo-1 ratio was  $8.1 \pm 0.4$  before, and increased to  $10.1 \pm 0.3$  after, break-in;  $n = 28$ ). In contrast, when 10 mM BAPTA buffer was included in the pipette solution,  $[Ca^{2+}]_i$  rapidly decreased (Fig. 2a; the indo-1

ratio was  $7.5 \pm 0.8$  before, and  $3.8 \pm 0.3$  after, break-in;  $n = 5$ ). To determine whether the observed  $[Ca^{2+}]_i$  elevation participated in the activation of the currents, we included  $Ca^{2+}$  buffers in the pipette solution. Figure 2b shows a dose-dependent block of current development by  $Ca^{2+}$  buffers. Thus,  $[Ca^{2+}]_i$  elevations participated in the activation of the NADPH oxidase under our experimental conditions.

We next investigated whether the addition of the chemotactic peptide f-MLP (*N*-formyl-methionyl-leucyl-phenylalanine) or the phorbol ester PMA (phorbol-12-myristate-13-acetate; these are known stimuli of NADPH oxidase activation) to the bath solution could elicit the currents. We added agonist to cells patch-clamped with 0.2 mM EGTA in the pipette solution (Fig. 2c). Spontaneous currents were relatively small under these conditions. f-MLP and PMA, but not the receptor antagonist boc-MLP (*N*-tert-butoxycarbonyl-methionyl-leucyl-phenylalanine), clearly activated the currents. Both the small spontaneous currents and the agonist-stimulated currents were blocked by addition of DPI, whereas  $Zn^{2+}$ , the  $H^+$ -channel inhibitor, did not prevent development of the currents.



**Figure 2** Activation of NADPH-oxidase currents by cytosolic  $[Ca^{2+}]_i$ , receptor agonists, and phorbol esters. Whole-cell currents and  $[Ca^{2+}]_i$  elevations were measured under different conditions of  $Ca^{2+}$  buffering. **a**,  $[Ca^{2+}]_i$  recording of cells patched using a non-buffered (control) or heavily  $Ca^{2+}$ -buffered (10 mM BAPTA) pipette solution. Data are shown as a ratio of indo-1 fluorescence emission. **b**, Peak whole-cell current densities in cells patched with the following pipette solutions (from left to right): nominally  $Ca^{2+}$ -free; 50  $\mu M$  indo-1 free acid; 0.2 mM EGTA; and 10 mM BAPTA. The number of analysed cells is indicated above the columns. **c**, Currents induced by the application of soluble stimuli to cells perfused with pipette solutions containing 0.2 mM EGTA. Conditions were, from top to bottom: control vehicle (dimethyl sulphoxide, DMSO, 0.1%); chemotactic peptide (f-MLP, 100 nM); receptor antagonist (boc-MLP, 1 mM); and phorbol ester (PMA, 5 nM) added in the presence or absence of  $Zn^{2+}$  (10  $\mu M$ ) in the bath solution. Arrowheads indicate the time of addition of the stimulus (2 min after break-in) and of the inhibitor DPI.



**Figure 3** Currents in eosinophils from patients with genetic defects. Whole-cell currents were activated using pipette solutions lacking  $Ca^{2+}$  buffers and containing either no NADPH (upper traces in **a**) or 8 mM NADPH (lower traces in **a**). **a**, Oxidase currents in eosinophils of a patient with G6PD, two patients with CGD and control eosinophils recorded under low-oxygen conditions. Arrowheads mark the onset of the whole-cell recording (break-in). **b**, Maximal current densities; conditions as in (**a**). Ctr, control. The number of analysed cells is indicated above the columns. **c**, Percentage of cells showing positive NBT staining and percentage of cells showing NADPH-oxidase currents in eosinophils from healthy volunteers (control, >10 different donors), from a CGD carrier, and from the two CGD patients.

Thus, the currents could be activated using pipette solutions without  $\text{Ca}^{2+}$  buffers (this situation elicits  $\text{Ca}^{2+}$ -mediated activation), or by soluble stimuli added to the bath solution.

We next studied eosinophils from patients with genetic defects related to superoxide production. Patients with glucose-6-phosphate-dehydrogenase deficiency (G6PD) show decreased  $\text{O}_2^-$  production because of decreased NADPH supply<sup>16</sup>. Without NADPH in the pipette solution, the eosinophils from G6PD patients showed

small currents. However, the currents were fully reconstituted when NADPH was included in the pipette solutions (Fig. 3a, b). We next studied two patients with chronic granulomatous disease (CGD). Without NADPH in the pipette, eosinophils from the two CGD patients did not show detectable currents. With NADPH in the pipette, the eosinophils of patient 2 displayed small currents ( $14 \pm 3\%$  of control), whereas the eosinophils of patient 1 did not display currents above background levels (Fig. 3a, b).

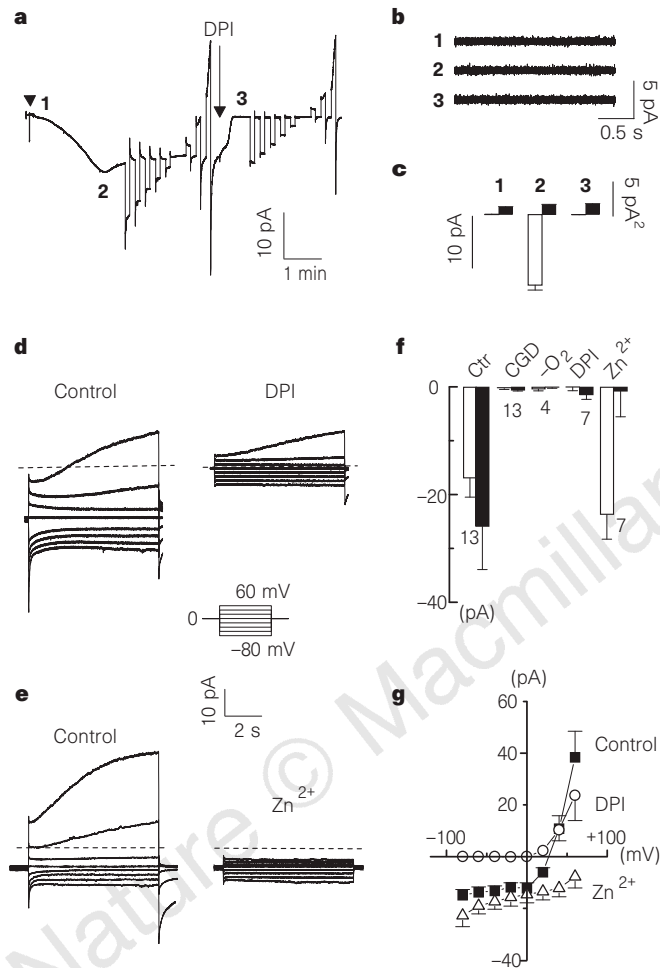
Electron-transfer currents should depend not only on the presence of the cytosolic electron donor (that is, NADPH) and a functional electron-transport chain (that is, NADPH oxidase), but also on the presence of an extracellular electron acceptor (that is, molecular oxygen). We therefore recorded current from cells in oxygen-depleted bath solutions<sup>17</sup>. No current developed without oxygen in the bath solution, even when NADPH was included in the pipette solution (Fig. 3a, b).

Heterozygous carriers of X-linked CGD do not have a clinically apparent disease; however, 50% of their eosinophils are unable to generate superoxide (because of X-chromosome inactivation). We therefore studied eosinophils from a carrier, and compared them with eosinophils from healthy volunteers and from the two CGD patients described above. NADPH-oxidase currents were found in most eosinophils from healthy controls, roughly half of the carrier eosinophils, and none of the CGD eosinophils (Fig. 3c). NBT precipitations were found in most control eosinophils, ~50% of the carrier eosinophils, and none of the CGD eosinophils (Fig. 3c).

The whole-cell patch-clamp technique provides information not only about the amplitude and kinetics of currents, but also about underlying unitary events. Indeed, the flux of ions through ion channels is discontinuous, because of the abrupt opening and closing of the unitary channels. This leads to characteristic whole-cell current fluctuations, which can be studied by noise analysis<sup>18</sup>. In contrast, electrogenic fluxes through a continuous transport mechanism are not expected to generate detectable noise. We therefore analysed the variance of the measured whole-cell currents at three different times: first, directly after break-in, that is, before relevant current activation; second, at the time of maximal current amplitude; and third, after current block by DPI (Fig. 4a). Inspection of the three current traces by eye revealed no visible differences in noise (Fig. 4b) and no difference in current variance (Fig. 4c). These results indicate the absence of detectable channel noise in our recordings, and are most compatible with a continuous flow of electrons through the NADPH oxidase.

We next studied the whole-cell currents during steps to various target voltages (Fig. 4d–g). The currents remained below the zero-current level at most voltages (dotted line in Fig. 4d, e). The current kinetics, however, were complex, because voltage-dependent activation and deactivation occurred during depolarizing and hyperpolarizing voltage steps, respectively (Fig. 4d, e). These voltage-dependent currents were identified as  $\text{H}^+$  currents on the basis of the following criteria: first, the kinetics of voltage-dependent activation and deactivation were similar to those of previously described  $\text{H}^+$  currents<sup>15,19</sup>; second, there was a reversible block by  $10 \mu\text{M}$   $\text{Zn}^{2+}$  (Fig. 4e) (not shown); and third, there was a reversal potential that followed the  $\text{H}^+$  equilibrium potential ( $E_{\text{H}^+}$ ; not shown).

The inward currents observed at the beginning of a depolarizing voltage step and at the end of a hyperpolarizing voltage step suggested the presence of voltage-independent current components. Consistent with the presence of a mixture of voltage-independent inward currents and voltage-activated outward  $\text{H}^+$  currents, the steady-state current–voltage relationship (Fig. 4g) was outwardly rectifying and reversed sign at around  $E_{\text{H}^+}$ . Addition of DPI completely blocked the inward currents and reduced the voltage-dependent outward currents (Fig. 4d, g). In contrast, addition of  $\text{Zn}^{2+}$  (Fig. 4e) led to a block of the voltage-dependent currents and revealed the presence of sustained inward currents,



**Figure 4** Noise analysis and voltage-dependence of the NADPH oxidase currents. Whole-cell currents were activated using pipette solutions lacking  $\text{Ca}^{2+}$  buffers and containing 8 mM NADPH. **a**, Continuous recording of a typical experiment. The current developed progressively after break-in (arrowhead) and was blocked by DPI ( $10 \mu\text{M}$ ). A target voltage protocol ( $-100$  mV to  $+60$  mV;  $20$  mV increments) was applied before and after DPI addition. **b**, Current sweeps ( $2$  s, filtered at  $2.5$  kHz, digitized at  $10$  kHz) taken immediately after break-in (**1**), during the peak current (**2**) and after DPI addition (**3**) (same cell as in **a**). **c**, Current sweeps as shown in **b** were analysed for mean current amplitude (white bars) and mean variance (black bars;  $n = 5$ ; ref. 23). **d**, **e**, Families of currents elicited by identical voltage protocols (inset; see also **a**) were obtained before and after addition of  $10 \mu\text{M}$  DPI, and before and after addition of  $10 \mu\text{M}$   $\text{Zn}^{2+}$ . Dotted lines indicate the zero current level. **f**, Voltage-independent currents (white bars) and voltage-dependent currents (black bars) represent sustained currents at  $0$  mV and voltage-dependent currents (black bars) represent deactivating inward currents seen after a voltage step from  $0$  mV to  $-80$  mV observed: under control conditions; in CGD eosinophils; in oxygen-depleted cells; after addition of  $10 \mu\text{M}$  DPI; and after addition of  $10 \mu\text{M}$   $\text{Zn}^{2+}$ . The number of analysed cells is indicated under the columns. **g**, Steady-state current–voltage relationship (determined at the end of the  $7$ -s voltage pulse) in control eosinophils or after addition of DPI or  $\text{Zn}^{2+}$  ( $n = 13$ ,  $7$ , and  $7$ , respectively). Residual currents in cells exposed to DPI and  $\text{Zn}^{2+}$  were used for leak subtraction.

with a linear current–voltage relationship over the whole range of voltages tested (Fig. 4g). The voltage-independent currents were identified as NADPH-oxidase currents on the basis of the following criteria: first, as expected for a unidirectional electron transporter, the currents were inward over the entire voltage range tested (–100 mV to +60 mV); second, the currents were blocked by DPI; third, the currents were resistant to the  $H^+$ -current inhibitor  $Zn^{2+}$ ; and fourth, the replacement of CsCl in the solutions with the impermeant ions *N*-methyl-D-glycine (NMDG<sup>+</sup>) and HEPES<sup>–</sup> did not modify the currents (not shown).

A puzzling observation of our study is the low threshold of voltage activation of the  $H^+$  conductance under experimental conditions that support NADPH-oxidase activity. This can be seen most clearly during a voltage step from 0 mV to –80 mV in control cells (Fig. 4d, e). The deactivating tail seen under these conditions shows that the  $H^+$  conductance was activated at 0 mV. Theoretically, however, the  $H^+$  conductance is expected to be activated at voltages of around +40 mV in the presence of a pH gradient of 0.5 pH units (ref. 20). DPI shifted the threshold of voltage activation of the  $H^+$  conductance towards the more positive voltages expected for our experimental conditions, as shown by the disappearance of the deactivating tail (Fig. 4d, f). No tail currents at –80 mV (that is, the sign of activated  $H^+$  conductance at 0 mV) were observed with other conditions that precluded NADPH-oxidase activation (for example, in CGD eosinophils or under conditions of no oxygen, Fig. 4f). Thus, there is a functional coupling between NADPH oxidase and the  $H^+$  conductance. Further studies will be necessary to elucidate the mechanism of coupling.

Our study is the first to demonstrate electron currents across the plasma membrane of a eukaryotic cell. These currents are generated by the phagocyte NADPH oxidase. The precise functions of the NADPH oxidase in host defence remain to be defined, however. The direct product of the NADPH-oxidase reaction, superoxide, is a relatively unreactive oxygen metabolite and its importance in microbial killing is unclear<sup>4</sup>. Our results indicate that the NADPH oxidase can generate large currents, even against major electrical gradients (Fig. 4g). This feature is unique compared with other electrogenic processes and might be—by itself—a biologically relevant function of the NADPH oxidase. Electron transport is a biologically relevant function of enzymes in, for example, bacteria, mitochondria and chloroplasts, where it generates protonmotive forces necessary for crucial cellular processes such as ATP generation or transport of protonated molecules<sup>1,21</sup>. Perhaps electron currents are involved in analogous vectorial transport in the phagocyte. □

## Methods

**Patients.** The G6PD patient was a 36-year-old white male with an erythrocyte glucose-6-phosphate dehydrogenase activity of 0.25 units per g haemoglobin (normal range: 3.5–5.5). CGD patient 1 was a 33-year-old white male with X-linked CGD. CGD patient 2 was a 33-year-old white female with gp47<sup>phox</sup>-negative CGD. The carrier was a 35-year-old healthy white female, whose brother had X-linked CGD<sup>22</sup>.

**Purification.** Purification of eosinophils and whole-cell patch-clamp studies<sup>15</sup>, noise analysis<sup>23</sup>,  $[Ca^{2+}]_i$  measurements<sup>24</sup>, NBT staining<sup>13</sup>, and oxygen depletion<sup>17</sup> were performed as described.

**Solutions.** If not indicated otherwise, the following solutions were used: bath solution: 75 mM CsCl, 50 mM CsOH, 50 mM HEPES/pH 7.1, 10 mM tetraethyl ammonium chloride (TEACl), 1 mM  $MgCl_2$ , and 0.1% glucose; pipette solution: 75 mM CsCl, 50 mM CsOH, 50 mM HEPES/pH 7.6, 10 mM TEACl, 1 mM  $MgCl_2$ , 1 mM  $MgATP$ , and 8 mM NADPH.

**Statistics.** Data are expressed as mean  $\pm$  s.e.m. For statistics, a Wilcoxon sign-rank test with a level of significance  $\alpha = 0.05$  was used.

Received 23 December 1997; accepted 2 February 1998.

1. Mitchell, P. Foundations of vectorial metabolism and osmochemistry. *Biosci. Rep.* **11**, 297–346 (1991).

2. Klebanoff, S. J. Oxygen metabolism and the toxic properties of phagocytes. *Ann. Intern. Med.* **93**, 480–489 (1980).
3. Clark, R. A. The human neutrophil respiratory burst oxidase. *J. Infect. Dis.* **161**, 1140–1147 (1990).
4. Segal, A. W. & Abo, A. The biochemical basis of the NADPH oxidase of phagocytes. *Trends Biochem. Sci.* **18**, 43–47 (1993).
5. Chanock, S. J., el Benna, J., Smith, R. M. & Babior, B. M. The respiratory burst oxidase. *J. Biol. Chem.* **269**, 24519–24522 (1994).
6. Clark, R. A. Genetic variation in chronic granulomatous disease. *Hosp. Pract.* **25**, 51–55 (1990).
7. Boxer, L. A. & Blackwood, R. A. Leukocyte disorders: quantitative and qualitative disorders of the neutrophil, part 1. *Pediatr. Rev.* **17**, 19–28 (1996).
8. Henderson, L. M., Chappell, J. B. & Jones, O. T. The superoxide-generating NADPH oxidase of human neutrophils is electrogenic and associated with a  $H^+$  channel. *Biochem. J.* **246**, 325–329 (1987).
9. Shult, P. A., Graziano, F. M., Wallow, I. H. & Busse, W. W. Comparison of superoxide generation and luminol-dependent chemiluminescence with eosinophils and neutrophils from normal individuals. *J. Lab. Clin. Med.* **106**, 638–645 (1985).
10. Gallin, E. K. & McKinney, L. C. in *Current Topics in Membranes and Transport*, Vol. 35: *Mechanisms of Leukocyte Activation* (eds Kleinzeller, A., Grinstein, S. & Rotstein, O. D.) 127–152 (Academic, San Diego, 1990).
11. Nanda, A. & Grinstein, S. The membrane potential of resting and activated neutrophils: determinants and significance. *Cell Physiol. Biochem.* **1**, 65–75 (1991).
12. Demareux, N., Schrenzel, J., Jaconi, M. E., Lew, D. P. & Krause, K.-H. Proton channels, plasma membrane potential, and respiratory burst in human neutrophils. *Eur. J. Haematol.* **51**, 309–312 (1993).
13. Newburger, P. E. et al. Parental diagnosis of chronic granulomatous disease. *N. Engl. J. Med.* **300**, 178–181 (1979).
14. Cross, A. R. & Jones, O. T. The effect of the inhibitor diphenylene iodonium on the superoxide-generating system of neutrophils. Specific labelling of a component polypeptide of the oxidase. *Biochem. J.* **237**, 111–116 (1986).
15. Schrenzel, J., Lew, D. P. & Krause, K. H. Proton currents in human eosinophils. *Am. J. Physiol. Cell Physiol.* **271**, C1861–C1871 (1996).
16. Vives Corrons, J. L. et al. Severe-glucose-6-phosphate dehydrogenase (G6PD) deficiency associated with chronic hemolytic anemia, granulocyte dysfunction, and increased susceptibility to infections: description of a new molecular variant (G6PD Barcelona). *Blood* **59**, 428–434 (1982).
17. Benesch, R. E. & Benesch, R. Enzymatic removal of oxygen for polarography and related methods. *Science* **118**, 447–448 (1953).
18. Hille, B. *Ionic Channels of Excitable Membranes* (Sinauer Associates, Sunderland, MA, 1992).
19. Gordienko, D. V. et al. Voltage-activated proton current in eosinophils from human blood. *J. Physiol.* **496**, 299–316 (1996).
20. Cherny, V. V., Markin, V. S. & De Coursey, T. E. The voltage-activated hydrogen ion conductance in a rat alveolar epithelial cells is determined by the pH gradient. *J. Gen. Physiol.* **105**, 861–896 (1995).
21. Saier, M. H. Peter Mitchell and his chemiosmotic theories. *ASM News* **63**, 13–21 (1997).
22. Zimmerli, W., Lew, D. P., Suter, S., Wyss, M. & Waldvogel, F. A. *In vitro* efficacy of several antibiotics against intracellular *S. aureus* in chronic granulomatous disease. *Helv. Paediatr. Acta* **38**, 51–61 (1983).
23. Varnai, P. et al. Highly cooperative  $Ca^{2+}$  activation of intermediate conductance  $K^+$  channels in HL-60 granulocytes. *J. Physiol. (Lond.)* **472**, 373–390 (1993).
24. Demareux, N. et al. Regulation  $Ca^{2+}$  influx in myeloid cells: role of plasma membrane potential, inositol phosphates, cytosolic free  $[Ca^{2+}]$ , and filling state of intracellular  $Ca^{2+}$  stores. *J. Clin. Invest.* **90**, 830–839 (1992).

**Acknowledgements.** We thank E. Huggler for technical assistance; L. Bernheim, S. Rawlings, and M. Rossier for discussions; W. Zimmerli, R. Seger and P. Beris for help with the patients; and N. Mensi for the G6PD determination. This work was supported by grants from the Swiss National Foundation (to J.S., K.-H.K., N.D., B.B.), the Janggen Poehn Foundation (J.S.), the Max Cloetta Foundation (N.D.), FEBS (B.B.), and the Swedish Medical Research Grant (L.S.).

Correspondence and requests for materials should be addressed to K.-H.K. (e-mail: kkrause@cmu.unige.ch).

## correction

# A late Neanderthal associated with Upper Palaeolithic artefacts

Jean-Jacques Hublin, Fred Spoor, Marc Braun, Frans Zonneveld & Silvana Condemi

*Nature* **381**, 224–226 (1996)

One of the important morphometric variables assessed in the comparisons between the bony labyrinths of Neanderthals and modern humans is the sagittal labyrinthine index. This index expresses what percentage of the posterior semicircular canal is situated inferiorly to the plane of the lateral semicircular canal. Its formula, as given in the legend to Fig. 2 of the above Letter, is incorrect and should be  $i/(s + i) \times 100$ . □



# Instrumental medley

Featured in this section are analytical instruments and accessories, such as an ICP-HEX-MS and software for searching spectral libraries, and systems for scanning electron microscopy.

## HyperJet GC-MS

From HD Technologies

*A bench-top GC-MS system that uses a time-of-flight mass analyser*

This system is equipped with a supersonic molecular-beam (SMB) interface and a dual-ion source. The system's interface facilitates on-the-fly coupling with any GC column flow rate. HyperJet also includes two ionization sources: an electron ionization (EI) source and a hyperthermal surface ionization (HSI) source. The EI source ionizes molecules within the supersonic beam as it passes through the source. This is said to reduce the probability of fragmentation, enhancing the molecular ions produced by many compounds. This mass spectrum can be used in library searches. Furthermore, the ions formed also retain a component of velocity from the expansion, allowing thermal background ions to be filtered out on the basis of energy, reducing this source of peak tailing. With the HSI source, ionization occurs as a result of collision of the molecular beam with a heated rhenium foil. HSI sensitivities are stated to be up to 100 times more sensitive by HSI than EI. The selectability of ionization means that many compounds can be analysed from dirty matrices, including aromatics directly from crude oils and drugs directly from urine. Both ionization sources provide different fingerprint spectra.

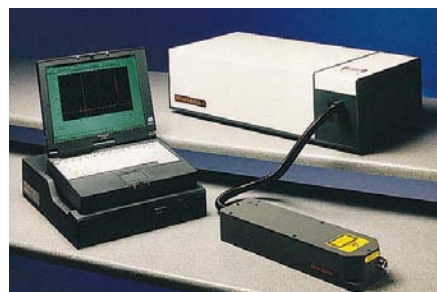
**Reader Enquiry No. 100**

## System 100

From Renishaw

*A compact Raman spectrometer for process control and general research applications*

The instrument uses dedicated fibre-optic probes as sampling stations, allowing sampling up to 100 metres from the spectrometer. Raman spectra can be acquired from samples whose sizes range from 1  $\mu\text{m}$  to more than 150 mm, with no sample preparation required. Each instrument consists of a control unit, a laser, a computer and



Renishaw's Raman probes have 100 m umbilicals.

up to four sampling stations. This allows the computer, control unit and laser to be isolated from the probe (this is useful for installations in cleanroom production facilities or when studying samples located in hazardous environments). The sampling stations are connected to the control unit by armoured fibre-optic cables up to 100 m in length.

**Reader Enquiry No. 101**

## Platform ICP

From Micromass

*The company's first instrument to use the technique of ICP-HEX-MS*

According to the manufacturer, this technique can be plagued by interference by the most important elements: K, Ca, Cr, Mn, Fe, As and Se. Such interference can result from the formation of molecular species in the plasma; combinations of matrix elements, such as  $^{40}\text{Ar}^{12}\text{C}^+$  on  $^{52}\text{Cr}^+$ ,  $^{40}\text{Ar}^{14}\text{N}^+$  on  $^{54}\text{Fe}^+$ ,  $^{40}\text{Ar}^{16}\text{O}^+$  on  $^{56}\text{Fe}^+$ ,  $^{40}\text{Ar}^{2+}$  on  $^{80}\text{Se}^+$  and  $^{40}\text{Ar}^{35}\text{Cl}^+$  on  $^{75}\text{As}^+$ . The instrument incorporates HEX declustering, utilizing a procedure that involves a hexapole ion lens, located behind the skimmer cone, which is surrounded by a gas cell. By admitting small amounts of helium, the pressure over the length of the hexapole is increased, breaking up molecular species before they reach the mass analyser. Detection limits for most elements are in the p.p.t. range. The instrument has a wide dynamic range photomultiplier detector that gives a stated eight orders of magnitude without having to change the detection mode.

**Reader Enquiry No. 102**

## Avance 700 spectrometer

From Bruker

*A new compact 700 Ultrashield magnet saves space in the laboratory for NMR instruments*

Featuring a compact, true ultra-high-field 700 MHz system for routine NMR operation, which is said to require less space than a traditional 500 MHz magnet. The technology is said to provide excellent field homogeneity, for high resolution and line shape. The magnet cryostat has been optimized to allow easy handling for cryogen refills, and the magnet size requires less than 3.65-m (12 foot) ceiling clearance for installation and operation. This instrument should allow those who do not have the space to accommodate a 750 MHz or 800 MHz system to have an instrument on site without resorting to costly site renovations.

**Reader Enquiry No. 103**



A massive spectacle: VG Axiom for ICP-MS.

## VG Axiom

From VG Elemental

*A third-generation high-resolution ICP-MS*

This instrument has an optional multicollector ion detection system in a compact housing with a high-power, magnetic-sector double-focusing spectrometer, a modern ICP ion source and transfer optics. The multicollector Axiom MC is said to perform simultaneous, accurate and precise isotope ratio measurements. The detector array is based on the VG Plasma 54 Multicollector ICP-MS. Both the single (Axiom SC) and multicollector instruments couple high ion transmission at all resolutions (>4 per cent at 10,000 RP), with stated background count rates of less than 0.2 counts per second, ensuring unparalleled detection limits on a routine basis. The resolution can be varied between 400 and >10,000 RP at any user-defined resolution.

**Reader Enquiry No. 104**

## MC-2000

From Ocean Optics

*A rack-mount system for multi-channel spectroscopy*

This is said to be a sturdy rack-mounted unit that accommodates up to eight individual S2000 series spectrometer channels for ultraviolet, visible and shortwave NIR spectroscopy (200–1,100 nm). All eight spectrometer channels operate from a single ISA-bus analog-to-digital (A/D) conversion card, making it well suited for process control applications and experiments requiring an expanded wavelength range, multiple-sample analysis and/or reference monitoring. Also, the rack mount can be configured to accept one or more spectrometer channels and one or more other analytical devices. With a compact optical bench coupled to a 2,048-element, linear CCD-array detector, the system should prove useful for fluorescence and other applications that demand high detector sensitivity.

**Reader Enquiry No. 105**

## Spectral ID

From Galactic

*Software that searches multiple 32-bit spectral libraries*

Allowing different ranges and resolutions to be searched, this software also permits users to save their favourite libraries for future searches, as well as defining custom libraries made up of virtually any optical data type at any spectral range or resolution. The 32-bit architecture takes advantage of Windows95 and 98, and NT 4.0 and 5.0 operating systems. Moreover, the system is said to reach search speeds of over 4,000 compounds per second on full spectrum searches and virtually instantaneous peak and text searches. Even using the region mask, a feature that allows users to include or exclude an unlimited number of spectral regions from a search, a feature that does not affect search speeds. The system offers seven full-spectrum search algorithms, including the accurate correlation and first derivative correlation, two methods of peak searching and flexible text query methods.

**Reader Enquiry No. 106**

## Eagle $\mu$ -Probe

From Edax

*A new generation of X-ray spectrometers for high sensitivity elemental analysis*

This compact, table-top analyser utilizes the latest in X-ray microbeam technology to provide fast and accurate elemental microanalysis. It uses X-ray capillary optics, which are said to allow the X-ray spot size to be concentrated to 100  $\mu\text{m}$  or less. The high intensities obtained with the optimized optics provide simultaneous non-destructive, multi-element analysis of a wide variety of sample types: solids, liquids and powders. The Eagle  $\mu$ -Probe is designed to handle material classification, qualitative and quantitative analysis of irregularly shaped, regularly structured or large solid samples, as well as plating analysis, layer analysis, and more.

**Reader Enquiry No. 107**

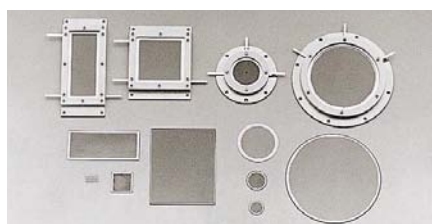
## Microchannel systems

### Microchannel plates

From Hamamatsu

*These plates, or MCPs, aid in the detection of ions in mass spectroscopy*

The plates feature compactness, light weight, the good timing properties that come with short length, high gain and pulse-height distribution (PHD). They consist of a two-dimensional array of millions of very small diameter glass capillaries fused together and sliced in the shape of a thin disc. These capillaries act as independent electron multipliers. In addition to previously available circular MCPs, Hamamatsu offers seven types of



Hamamatsu's circular and rectangular MCPs.

rectangular MCP's. These are designed to alleviate dimensional limitations encountered when building the MCPs into instruments as compared with circular types.

**Reader Enquiry No. 108**

## Amplifier/discriminator

From EG&G Ortec

*A new 1-GHz amplifier and timing discriminator for ultra-fast detectors*

The highly compact model 9327 is designed to deliver optimum time resolution for microchannel plate detectors, microchannel plate photomultiplier tubes, fast photodiodes and fast photomultiplier tubes. It can also be used for single-photon or single-ion counting with those detectors. Input pulse amplitudes from 0 to  $-30\text{ mV}$  or 0 to  $-150\text{ mV}$  are accommodated, says EG&G. Moreover, an over-range LED allows accurate, fine-gain adjustment without the use of an oscilloscope. The compact package allows the unit to be located close to the detector, thus minimizing ground loops and pick-up noise. This model delivers a timing jitter  $< 20\text{ ps}$  at half scale and walk of less than  $\pm 40\text{ ps}$  over the top 90 per cent of full scale.

**Reader Enquiry No. 109**

## SEM systems

### NetSEM

From Leo Electron Microscopy

*The LEO 400 SEM series may now be interfaced to the Internet*

NetSEM makes it possible for users to log onto this scanning electron microscopy (SEM) system from remote locations and not only read all instrument conditions, but also to control them in real time. All essential instrument parameters are available for control, and images are compressed with the latest techniques for transmission. The basic version of NetSEM is available free of charge and the software may be downloaded from the company's website. It is especially useful for allowing system supervisors to monitor their microscopes, or assist users, from remote locations. The program requires only an Internet browser and can therefore be used from an office within the same site, from another country or even using a portable PC with mobile communication. It allows multiple users to log on simultaneously in a teaching envi-

ronment or conference microscopy. (For more information on telemicroscopy see *Nature* 391, 613-614; 1998).

**Reader Enquiry No. 110**

## SEM vacuums

From Electron Microscopy Sciences

*A complete line of vacuum equipment*

This line of equipment includes stand-alone glow discharge units, etch facilities, film thickness monitors, carbon-coating units, sputter coating units, turbo-sputter and turbo-carbon coaters, critical-point driers, freeze dryers, turbo evaporators, plasma ashers and cryogenic preparation systems. All of the equipment in the line is fully automatic and comes with an unconditional two-year warranty on all parts and labour, says the manufacturer. Some of the standard features bundled with this equipment include built-in rotary and tilt stages, built-in multispecimen holders, built-in penning head and gauges, to name a few. The equipment is designed in such a way to accommodate multiple accessories and attachments.

**Reader Enquiry No. 111**

*These notes are compiled by Brendan Horton from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.*

## ADVERTISEMENTS

**Custom antibody production**

**David's Biotechnologie** Planning - production - after sales support

**chicken egg yolk antibodies**  
- fast - cost effective - very high yield -

**rabbit antibodies**

**monoclonal antibodies**  
- development and production -

**NEW - recombinant antigen service -**

Ask for our catalog: David's Biotechnologie  
Karl Anselm Str.1A, D-93051 Regensburg  
Germany, email: David's Bio@T-online.de  
Phone: +49(0)171/4900308 Fax/Phone: +49(0)941/948228

## READER ENQUIRY NO. 27

### ANTIBODIES

For Research & Routine Histology

Inflammation - Angiogenesis - Tumors  
Leukocytes - Endothelial Cells  
Pathological Markers

Custom Monoclonals and Polyclonals

BMA Biomedicals Ltd.  
(800) 634-3183 (US)  
Int'l: +41 61 811 62 22

## READER ENQUIRY NO. 28